

# 九州工業大学学術機関リポジトリ



Title	予測符号化のためのニューロンネットワークの学習精度向上に関する研究
Author(s)	小早川, 俊祐
Issue Date	2010-09
URL	<a href="http://hdl.handle.net/10228/5475">http://hdl.handle.net/10228/5475</a>
Rights	

# 博士学位論文

## 予測符号化のための ニューロンネットワークの 学習精度向上に関する研究

平成22年9月

九州工業大学大学院生命体工学研究科

小早川 俊祐

# 目次

## 第1章 序論

1.1 研究の背景	1
1.1.1 予測器と予測符号化の研究動向	1
1.1.2 予測符号化のためのニューロンネットワークの問題点	4
1.2 研究の目的及び論文の構成と概要	5
参考文献	8

## 第2章 入力信号数増加によるニューロンネットワークの学習精度向上

2.1 緒言	10
2.2 計算機シミュレーション	10
2.2.1 方法	10
2.2.2 結果	12
2.3 考察	16
2.4 結言	17
参考文献	18

## 第3章 並列化によるニューロンネットワークの学習精度向上

3.1 緒言	19
3.2 並列型ニューロンネットワーク	19
3.2.1 入出力特性	19
3.2.2 学習則	21
3.3 計算機シミュレーション	22
3.3.1 方法	22
3.3.2 結果	24
3.4 考察	34
3.5 結言	35
参考文献	36

## 第4章 ボルテラフィルタ内蔵によるニューロンネットワークの学習精度向上

4.1 緒言	37
4.2 ボルテラニューロンネットワーク	38
4.2.1 ボルテラニューロン	38

4.2.2	入出力特性	41
4.2.3	学習則	42
4.3	計算機シミュレーション	44
4.3.1	方法	44
4.3.2	結果	47
4.4	考察	51
4.5	結言	52
	参考文献	54

## 第5章 パラメータの時変化によるボルテラニューロンネットワークの学習精度向上

5.1	緒言	58
5.2	時変ボルテラニューロンネットワーク	59
5.2.1	入出力特性	59
5.2.2	学習則	61
5.3	学習精度の評価法	64
5.4	考察	65
5.5	結言	66
	参考文献	68

## 第6章 誤差収束の原理

6.1	緒言	69
6.2	原理	70
6.3	工学的実現	71
6.4	応用例	76
6.5	考察	78
6.6	結言	79
	参考文献	80

## 第7章 誤差収束によるニューロンネットワークの学習精度向上

7.1	緒言	81
7.2	誤差収束型ニューロンネットワーク予測器	81
7.3	計算機シミュレーション	88
7.3.1	方法	88
7.3.2	結果	92

7.4 考察	101
7.5 結言	103
参考文献	105
<b>第8章 誤差収束型ニューロンネットワーク予測器を用いた予測符号化</b>	
8.1 緒言	106
8.2 独立量子化型学習予測符号化	107
8.2.1 独立量子化	107
8.2.2 学習機能の導入	108
8.2.3 誤差の低減化	109
8.3 計算機シミュレーション	110
8.3.1 方法	110
8.3.2 結果	112
8.4 考察	117
8.5 結言	118
参考文献	120
<b>第9章 結論</b>	
9.1 本論文の要約	121
9.2 今後の課題	122
謝辞	123
本論文に関連する研究業績リスト	124

# 第1章 序論

## 1.1 研究の背景

### 1.1.1 予測器と予測符号化

予測とは、一般に過去から現在までの情報を基にして、事前に未来の現象を推し測ることである。これを予知、予見ともいう。生物は、目的のために予測を行い、その予測結果を用いて適切な行動を行おうとする。これを応用したものに予見制御がある。古来より人類も、過去から現在までの情報や経験則に基づいて、生存のために、様々な予測を行い生活に利用してきた。例えば、危険予知、地震予知、天気予報、経済予測などがある。これらを実現する予測方法には、統計をはじめとする、機械学習、データマイニング、カオス理論、シミュレーション、信号処理などを用いたものがある。中でも、シミュレーションと信号処理などには、時系列信号を予測するための予測器を用いる手法がある。この予測器は、時系列信号の特性に合わせて適切な予測法が適用される。

時系列信号の予測には、未知の予測値が未来値の前向き予測、過去値の後ろ向き予測がある。また、予測法には、予測対象の信号を線形として取り扱う線形予測法と非線形として取り扱う非線形予測法がある。これらの予測法は、その予測を行うために利用するデータの時間区間によって、局所的、中域的、大域的の予測に区分できる。各予測の特徴は、局所的予測は予測時刻付近のデータによる予測、中域的予測は予測を行うための最適区間のデータによる予測、大域的予測は過去から予測時刻直前、または、現在から予測時刻直後までのすべてのデータによる予測といえる。一般に時系列信号の予測は、大域的予測が行われる。さらに、非線形予測法は、予測対象の信号を微小区間に分割し、それぞれの区間を線形近似して取り扱う線形予測と任意の区間を非線形のまま取り扱う非線形予測を用いるものに区分される。予測器のパラメータの時不変、時変によっても区分される。時変予測器ならば、適応予測となる。また、予測対象の時系列信号に対して、予測器を専用と汎用のものに区分できる。

実際の時系列信号は、そのほとんどが非線形であるが、微小区間では線形で近似できるため、専用の線形予測器の利用も根強い。しかし、これも次第に非線形予測器になってきている。汎用の線形予測器は、主に線形時系列信号、または、高精度が要求されない予測に使用されている。線形予測法には、統計的手法が主に用いられ、単純平均、移動平均、簡易指数平均、重

回帰, 移動平均自己回帰など多くある. 一般的な予測器には, 自己回帰を適用するものが多い. この予測器は, 有限インパルス応答フィルタや1次ボルテラフィルタを応用したものと表現することもある. フィルタには, ウィーナーフィルタなどがあり, 線形動的システムに対しては, 主にカルマンフィルタを適用している. 確率的手法としては, 隠れマルコフモデルなどがあり, 最近では, カルマンフィルタに代わって, パーティクルフィルタが台頭しつつある. その他代表的なものに, 予測と量子化を同時に行う予測量子化器を用いる差分パルス符号変調 (Differential Pulse Code Modulation, 以下DPCMと称する) や, これよりさらに圧縮率を上げるために, 量子化幅を適応的に変化させる適応的DPCMなどがある.

一方, 非線形予測法を用いた予測器には, 非常に様々なものがある. 予測器へ適用する専用非線形モデルには, 航空機の縦系非線形運動方程式, 大気変動を表わすローレンツ方程式, 経済予測に用いられる自己回帰条件付不均一分散モデル (Autoregressive Conditional Heteroskedasticity: ARCH) や一般化自己回帰条件付不均一分散モデル (Generalized Autoregressive Conditional Heteroskedasticity: GARCH) などをはじめとし, 非常に多くある. 汎用モデルを用いる予測器に適用する線形予測としては, 線形予測法で用いる手法をそのまま局所線形予測に適用する方法をはじめ, ヤコビ行列推定, 動径基底関数 (Radial Basis Function, 以下RBFと称する) などがある. 決定論的[1]には, ローレンツ類推法, ローレンツ類推法とブートストラップ法の融合法, SugiharaとMayの方法, 池口と合原による改良類推法, FarmerとSidorowitchによるヤコビ行列推定法, Lebaronの方法, 五百旗頭らによる局所ファジィ再構成法, Meesによるボノロイ分割法などの局所線形予測がある. 最近では, 最新データの取舍を選択しながらの線形近似[2]もある. 非線形予測には, 数学的手法では, 平滑化スプライン, カーネル法[3]など, フィルタでは, 2次ボルテラフィルタ (Second-Order Volterra Filter, 以下2VFと称する) など, ニューロンネットワークでは, 多層パーセプトロン, RBF組込み型多層パーセプトロン, RBFネットワーク (Radial Basis Function Network [4], 以下RBFNと称する), RBFNとDPCMの組合せ, リカレントニューロンネットワーク (Recurrent Neuron Network, 以下RNNと称する), 入力遅れニューロンネットワーク (Input-Delay Neuron Network, 以下IDNNと称する), 時間遅れニューロンネットワーク (Time-Delay Neuron Network [5], 以下TDNNと称する), ボルテラニューロンネットワーク (Volterra Neuron Network, 以下VNNと称する) などがある. 以上列記したように, 予測器に適用する予測モデルが非常に多くあり, その選択を行わなければならない. モデル選択基準として, 統計学では, 最終予測誤差, 赤池情報量規準, ベイズ情報量規準, 交差検定, 情報理論では, 最小記述長などがある.

その他、予測器に関することとしては、予測器の構成や数値処理を工夫して予測精度を上げるアンサンブル予測などがあり、天気予報も近年この手法を用いるようになった。また、計算機の処理速度向上のために条件分岐命令の分岐予測のための予測器、すなわち、分岐予測器やその予測精度を向上させたハイブリッド予測器がある。最先端の分岐予測には、ニューロン分岐予測が採用されており、パーセプトロン分岐予測器がある。

時系列信号処理用予測器は、予測符号化で信号圧縮のために用いられる。予測符号化は、主に線形と非線形に区分され非可逆圧縮を行う。線形予測法を用いるものを線形予測符号化（Linear Predictive Coding, 以下LPCと称する）と称する。非線形予測法を用いるものを非線形予測符号化（Nonlinear Predictive Coding, 以下NLPCと称する）と称する。予測符号化は、はじめ線形予測法を用いていた。しかし、予測対象の時系列信号のほとんどが非線形であることから、信号圧縮率向上のために、1990年代頃になって、ようやく非線形予測法を用いたNLPCの研究が行われるようになった。線形予測器という言葉が最初に用いられたのは、1949年アメリカ合衆国のN. Wienerによる著書[6]の中である。音声信号へ線形予測法を最初に応用したのは、1966年日本電信電話公社の斉藤と板倉が最尤法を使った自動音素識別法による音声符号化である。1967年J. P. Burgが地震波解析に関して、時系列に線形予測モデルを当てはめる最大エントロピー法[7]を使った手法を提案した。1969年斉藤と板倉は、さらに部分相関法を発表した。同年G. J. Cullerは、リアルタイム音声符号化を発表している。また、1967年から1970年にかけて、B. S. AtalとR. W. SchaferやM. R. Schroederによって、適応予測符号化や線形予測音声符号化などの成果がもたらされた[8]。このようにして、LPCの基礎が確立され、次第にオンラインシステムに適合したLPCの改良が行われるようになった。

一方、非線形予測法は、従来から多くあるものの、NLPCの研究の歴史は浅い。このため、従来のNLPCを改善することを目的とする研究が少なく、LPCを改善する目的だけの様々なNLPCが提案されてきた。ニューロンネットワークは、NLPCの研究の初め頃からすでに非線形予測器へ用いられている。1993年L. Wuら[9]が、RNNを用いて音声符号化を行い、線形予測音声符号化より能力が改善されることを示した。同年R. J. P. deFigueiredo[10]は、フォック空間内における非線形予測符号化のためのニューロンネットワークモデルを提案した。1994年J. Thyssenら[11]は、音声について2VFとIDNNを用いた2種類のNLPCとLPCの比較をし、音声の非線形性の証明を行った。1995年F. Dz'az-de-Mariaら[12]は、音声符号化のために、モデルパラメータ数と計算負荷の低減ができる特徴を備えたRBFを用いて非線形予測を行った。同年S. W. Panら[13]は、LPCと複数の線形項の積に比例する交差項を付加したNLPCを含む音声圧縮のための方



法とシステムの特許出願を行った。1998年A. Fukunagaら[14]は、可逆画像圧縮を行うために、遺伝的プログラミングを用いた非線形予測器で、改めて量子化を行わない画像データ専用の予測符号化を提案した。2003年M. Vandana[15]は、音声のM配列予測符号化を提案した。2006年M. Faundez-Zanuy [16]は、アンサンブル法ニューロンネットワークを用いた音声符号化を提案した。2007年N. Stefanoskiら[17]は、非線形予測器を用いた時不変動的3次元メッシュの階層予測符号化を提案した。その他の予測符号化として、2009年太田ら[18]は、2値系列だけで処理を行う反辞書式のデータ処理と圧縮率改善のための反辞書式予測符号化を提案した。

圧縮率の高い高性能な予測符号化を行うためには、汎化精度が高い予測器が必要である。この要求に対応するために、前記NLPCの動向に見られるようなニューロンネットワークを用いた予測器が、他の非線形予測を用いた予測器に代わって台頭してきた。しかし、汎化精度は十分とはいえず、学習用信号に対しても誤差が生じるという問題が依然と残る。また、予測符号化の量子化過程で、量子化系列信号に誤差が含まれるという予測符号化の原理に関しても問題がある。

### 1.1.2 ニューロンネットワークの学習精度

ニューロンネットワークには、学習、汎化、実現の3つの能力が備わっている。しかし、これらの能力は、目的に対しての課題を完全に遂行する保障をしない。ニューロンネットワークは、各層への入力信号と教師信号の間に関数関係がなければ、学習できない。したがって、入力信号と教師信号との間に関数関係がある成分とない成分が混在する場合、入力信号と教師信号との間に関数関係がない成分は学習できない。さらに、ニューロンネットワークの学習精度が低い場合、入力信号と教師信号との間に関数関係がある成分は、完全に学習できない。この関数関係がある成分を完全に学習するためには、ニューロンネットワークの学習精度を向上させることが考えられる。このように、ニューロンネットワークの基本性能を決定付ける学習精度に関する研究は、重要な位置を占める。また、出力精度が要求される応用分野では、ニューロンネットワークの性能は、まだ不十分であり、実際に用いられることが少ない。したがって、従来からニューロンやニューロンネットワークの学習精度や構造を改善し、出力精度の向上が試みられてきた。これらを改善したものには、高次[19]、入力遅れ、RBF、ボルテラなどの各ニューロン、GMDH (The Group Method of Data Handling [20])、IIR (Infinite impulse response [21])、ピラミッド[22]、2次元拡散連想[23]、並列型ボルテラ[24]、介在素子付加型[25]などの

各ニューロンネットワークがある。しかし、これらを用いたニューロンネットワークは、入力信号と教師信号との間の関数関係に対して十分に学習を行っても、出力誤差が全くない完全学習を保証できないという問題がある。このため、ニューロンネットワークは、誤差の許容範囲の中で使用されているのが現状である。

## 1.2 研究の目的及び論文の構成と概要

本研究の目的は、前記問題点を解決するために、予測器に用いるニューロンネットワークの構成を工夫し、学習精度を高めた予測器で汎化精度を向上させ、圧縮率が改善された量子化原系列信号についての可逆圧縮方法としての新しい処理構成の予測符号化を提案することにある。さらに、学習用の時系列信号に対して、予測器が誤差の全くない完全な学習を成し遂げたときに得られる時系列信号の保存機能も実現させる。これにより、学習可能な高精度予測器が必要な航空宇宙機の知的予見飛行制御への応用へも期待できるものとなる。

本論文は、時系列信号に関して学習精度を高めたニューロンネットワーク予測器を用いた高い汎化精度の実現とそれを用いた予測符号化の改善に至るまでの過程を第1章序論、第2章入力信号数増加によるニューロンネットワークの学習精度向上、第3章並列化によるニューロンネットワークの学習精度向上、第4章ボルテラフィルタ内蔵によるニューロンネットワークの学習精度向上、第5章パラメータの時変化によるボルテラニューロンネットワークの学習精度向上、第6章誤差収束の原理、第7章誤差収束によるニューロンネットワークの学習精度向上、第8章誤差収束型ニューロンネットワーク予測器を用いた予測符号化、第9章結論の9つの章で論述している。

第1章では、研究の背景として、予測器と予測符号化の研究動向を述べ、ニューロンネットワークを用いた予測符号化への変遷とその問題点を説明している。そして、ニューロンネットワークの問題点と応用をかんがみながら、研究の目的及び本論文の構成と概要を述べている。これにより、研究の位置付けを確認する。

第2章では、BP ネットワーク（Back-propagation Network, 以下 BPN と称する）の入力信号を付加することで、学習能力にどのような影響があるのかを調べるために、入力信号数が異なる2つの非線形信号予測システムニューロンネットワークを用いて、それらの学習能力の評価を行った結果について述べている。そして、入力信号数による学習能力への影響を正しく認識する。

第3章では、2出力以上のBPNの出力間で生じる学習の干渉や各出力で用いる中間層の共有化実現の困難性の問題を指摘している。この問題によるBPNの能力低下を改善するためのニューロンネットワークとして、並列型ニューロンネットワーク (Parallel-Type Neuron Network、以下PNNと称する) を説明している。そして、PNNの問題改善に対する有効性を示すために、PNNのすべての出力の学習精度に関して、BPNに対する優位性が存在することを実証している。これにより、正しいニューロンネットワークの構成方法を認識する。

第4章では、心電図信号の非線形予測が抱えているQRS波の予測誤差問題に対する2次ボルテラニューロンネットワーク (Second-Order Volterra Neuron Network、以下2VNNと称する) の改善能力を検討している。具体的には、IDNNと2VNNによる正常洞調律心電図信号の非線形予測学習を評価し、2VNNの学習能力を示している。これにより、2VNNの学習能力のレベルを正しく認識する。

第5章では、学習計算機シミュレーション時間が比較的少なく、かつ、非線形予測能力がある1次ボルテラニューロンネットワーク (First-Order Volterra Neuron Network、以下1VNNと称する) を用いて、時変VNN (Time-Variant Volterra Neuron Network、以下TVVNNと称する) を構成し、不整脈心電図信号予測に関してのTVVNNの有効性を評価する方法を検討する。

第6章では、関数関係がある入出力を学習可能な近似関数とその近似関数によって生じる誤差を次々と別の学習可能な近似関数で表し、誤差を収束させる誤差収束の原理を提案する。また、この原理の工学的な実現方法として、量子化信号の誤差収束の原理を提案する。さらに、この原理を具現した誤差収束関数システム (Error-Convergence Function System、以下ECFSと称する) とそれを適用した誤差収束型ニューロンネットワーク (Error-Convergence Neuron Network、以下ECNNと称する)、誤差収束並列関数システム (Error-Convergence Parallel Function System、以下ECPFSと称する) とそれを適用した誤差収束並列型ニューロンネットワーク (Error-Convergence Parallel Neuron Network、以下ECPNNと称する)、誤差収束リカレントシステム (Error-Convergence Recurrent System、以下ECRSと称する)、誤差収束並列リカレントシステム (Error-Convergence Parallel Recurrent System、以下ECPRSと称する) を提案している。そして、それらによる関数システムの学習精度向上を理論的に説明している。これにより、従来のニューロンネットワークシステムの学習能力の限界に対する認識を改める。

第7章では、ECNNを予測符号化の予測器に適用するために、ECNNに用いられているニューロンネットワークの入力信号と教師信号の関数関係を強くするための再設計を施した誤差

収束型ニューロンネットワーク予測器 (Error-Convergence Neuron Network Predictor, 以下 ECNNP と称する) の高精度予測器としての実現性と時系列信号の保存機能を確認する。

第 8 章では, 汎化利用時の ECNNP の高段のニューロンネットワークの誤差の補償効果の確認, ECNNP の汎化能力の評価, 予測器に ECNNP を用いた予測符号化の改善に関して述べている。そして, 予測符号化のデータ圧縮率と再構成系列信号の精度を改善できる独立量子化型予測符号化 (Independent Quantization Predictive Coding, 以下 IQPC と称する), リアルタイムに ECNNP のニューロンネットワークを学習しながら予測誤差を低減する独立量子化型学習予測符号化 (Independent Quantization Learning Predictive Coding, 以下 IQLPC と称する), ニューロンネットワークの教師信号の符号にそのニューロンネットワークの出力信号の符号を合わせて誤差の低減を行う方法, IQLPC と Cabinet (以下 CAB と称する) と称するデータフォーマットを用いた 2 重圧縮の 4 つの考案に関して述べている。

第 9 章では, 本研究で得られた成果と今後の課題について述べている。

## 参考文献

- [1] 上野公彦, “数理水産科学のための時系列データの幾何学的解析入門-決定論的観点からの考察-,” 数理水産科学, vol.3, pp.103-155, 2005.
- [2] F. Kwasniok and L. A. Smith, “Real-time Construction of Optimized Predictors from Data Streams,” Physical Review Letters, vol.92, no.16, pp.164101-1-4, 2004.
- [3] SK. Lin, JZ. Qiao, GR. Wang, G. Zheng and J. Dong, “Nonlinear Time Series Prediction Modeling Based on Kernel Method,” Jisuanji Gongcheng / Computer Engineering, vol.33, no.17, pp.23-25, 2007.
- [4] T. Possio and F. Girosi, “Networks for approximation and learning,” Proc. of the IEEE, vol.78, no.9, pp.1481-1497, 1990.
- [5] K. J. Lang and G. E. Hinton, “A Time-Delay Neural Network Architecture for Speech Recognition,” Carnegie Mellon University Computer Science Technical Report, CMU-CS-88-152, pp.1-37, 1988.
- [6] N. Wiener, “Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications,” The MIT Press, pp.1-176, 1949.
- [7] 森下巖, 小畑秀文, “信号処理,” 計測自動制御学会, 初版第7刷, pp.145-148, 1999.
- [8] S.Haykin 著, 武部幹訳, “適応フィルタ入門,” 現代工学社, 第4版2刷, pp.21-22, 85-86, 2003.
- [9] L. Wu, M. Niranjan and F. Fallside, “Nonlinear Predictive Vector Quantization with Recurrent Neural Nets,” Neural Networks for Signal Processing [1993] III. Proceedings of the 1993 IEEE-SP Workshop, pp.372-381, 1993.
- [10] R. J. P. deFigueiredo, “A Neural Network Model for Nonlinear Predictive Coding in Fock Space,” 1993 IEEE International Symposium on Circuits and Systems, vol.4, pp.2164-2167, 1993.
- [11] J. Thyssen, H. Nielsen and S.D. Hansen, “Non-linear Short-term Prediction in Speech Coding,” 1994 IEEE International Conference on Acoustics, Speech, and Signal Processing, vol.1, pp.185-188, 1994.
- [12] F. Diaz-de-Maria and A. R. Figueiras-Vidal, “Nonlinear Prediction for Speech Coding Using Radial Basis Functions,” 1995 International Conference on Acoustics, Speech and Signal Processing, vol.1, pp.788-791, 1995.
- [13] S. W. Pan, SP. T. Wei and N. M. Labun, “Method and System for Compressing a Speech Signal Using Nonlinear Prediction,” US Patent 5696875, pp.1-8, 1997.

- [14] A. Fukunaga and A. Stechert, "Evolving Nonlinear Predictive Models for Lossless Image Compression with Genetic Programming," Proceedings of 3rd Annual Genetic Programming Conference, San Francisco, USA, pp.1-8, 1998.
- [15] M. Vandana, "M-ary Predictive Coding: A Nonlinear Model for Speech," Workshop on Spoken Language Processing, pp.59-66, 2003.
- [16] M. Faundez-Zanuy, "Speech Coding Through Adaptive Combined Nonlinear Prediction," Speech Communication, vol.48, issue 7, pp.838-847, 2006.
- [17] N. Stefanoski, P. Klie, X. Liu and J. Ostermann, "Layered Predictive Coding of Time-Consistent Dynamic 3D Meshes Using a Non-Linear Predictor," 2007 IEEE International Conference on Image Processing, vol.5, pp.109-112, 2007.
- [18] 太田隆博, 森田啓義, "多値アルファベットに対する反辞書を用いた心電図の予測符号化," 信学技報, MBE2009-44, pp.35-40, 2009.
- [19] C. L. Giles and T. Maxwell, "Learning, Invariance and Generalization in High Order Neural Networks," Applied Optics, vol.26, no.23, pp.4972-4978, 1987.
- [20] A. G. Ivakhnenko, "The Group Method of Data Handling-A Rival of the Method of Stochastic Approximation," Soviet Automatic Control, vol.13 c/c of Avtomatika, 1, 3, pp.43-55, 1968.
- [21] A. D. Back and A. C. Tsoi, FIR and IIR Synapses, "A New Neural Network Architecture for Time Series Modeling," Neural Computations, vol.3, pp.375-385, 1991.
- [22] 星野美保, 北村佐久, 増田健, 鈴木正樹, 趙晋輝, "RBF ネットワークの多層化とピラミッド型ネットワークの提案," 2000 年情報・システムソサイエティ大会講演論文集," p.28, 2000.
- [23] 木下智樹, 中村 清実, "二次元呈示位置自由度を持つ物体の形状と位置を同時認識する二次元拡散連想型ニューラルネット," 信学技報, NC97-166, vol.97, no.623-624, pp.209-216, 1998.
- [24] 小早川俊祐, 横井博一, "航空機ピッチ姿勢制御のためのボルテラフィルタ内蔵型ニューラルネットワーク," 平成 17 年度電気関係学会九州支部連合大会講演論文集," p.429, 2005.
- [25] 横井博一, 木本智幸, "介在素子を付加した階層型ニューラルネットワーク," バイオメディカル・ファジィ・システム学会誌, vol.1, no.1, pp.87-97, 1999.

## 第2章 入力信号数増加によるニューロンネットワークの学習精度向上

### 2.1 緒言

一般に、時系列信号を1つの入力信号と1つの教師信号で学習するBPN[1]は、1つの入力信号値に対して、2つ以上の教師信号値が対応する場合がある。このような状態で、BPNを学習させることはできない。この問題を解決するために、時間遅れ[2]、入力遅れ[3]、ボルテラ[4],[5]、IIR[6]などの各ニューロンネットワークを用いる手法、BPNへ教師信号と相関のある入力信号を付加する手法などを用いて、1つの入力信号値に対して、2つ以上の教師信号値が対応する確率を低減させている。しかし、他の応用分野では、BPNの入力信号を減らすことで、BPNの演算時間を短くする研究[7]も行われている。

本章では、BPNへ教師信号と相関のある入力信号を付加することで、学習精度にどのような影響があるのかを調べるために、入力信号数が異なる2つの非線形信号予測システムニューロンネットワークを用いて、それらの学習精度の評価を行った結果について述べる。第7章では、その結果を改善された予測符号化へ用いるニューロンネットワーク予測器の入力層の設計手法として用いる。

### 2.2 計算機シミュレーション

#### 2.2.1 方法

図2.1のBPNの時刻 $\tau$ の入力に対する時刻 $\tau+1$ の予測を用いて、2入力1出力と5入力1出力の2つの非線形信号予測システムを構成する。ここで、 $\tau$ は離散時間を示す整数である。そして、図2.2に示す離散時間間隔20[ms]の学習用の入力信号と教師信号、各200個を用いて、両システムの学習を行う。信号 $x_1 \sim x_4$ と $x_5$ は、それぞれ非線形プラントの運動方程式から得られた信号とそのプラントの制御信号にゲインチューニングを施した信号である。ゲインチューニングは、非線形プラントシミュレータから得られた信号を、出力を決定しやすい入力信号へ変換するためや、学習が効率的に行われるように入力の総和が出力関数の効率的な運用範囲内の値となるように調節するためのもので、各信号値を-1~1の範囲に設定する。ここで、2入力の

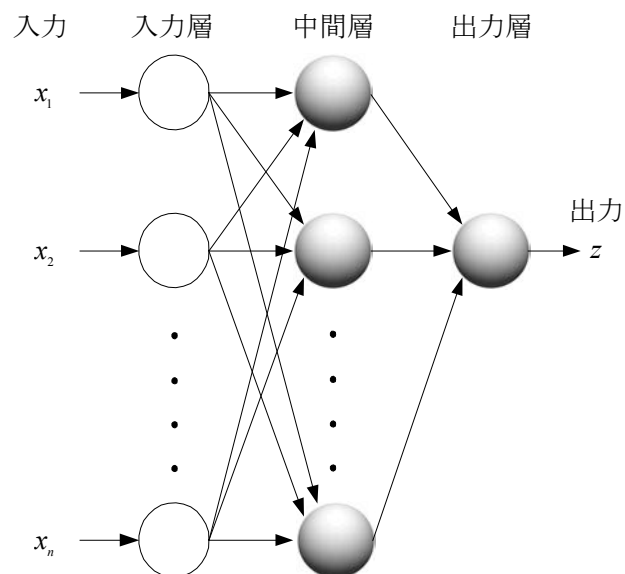


図 2.1 1 出力の BP ネットワーク

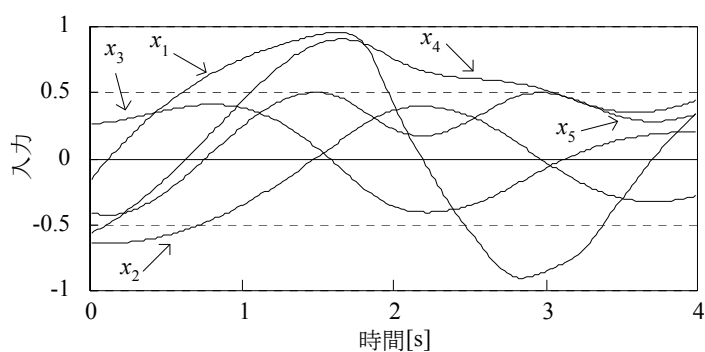


図 2.2 学習用非線形プラント信号

表 2.1 学習精度評価計算機シミュレーション条件

BPNの種類			2入力		5入力	
条件項目						
学習則			誤差逆伝搬則			
初期値		結合荷重		-0.3～0.3		
		しきい値		-0.3～0.3		
中間層素子数			範囲		2～30	
			間隔		2	
学習強化係数	勾配法項	粗探索	範囲		$10^{-4} \sim 1$	
			間隔		10倍	
		細探索	範囲		$0.2 \alpha_{min} \sim 0.9 \alpha_{min}$	
			間隔		$0.1 \alpha_{min} \sim \alpha_{min}$	
	慣性項				0	
	学習サイクル			30,000		
施行回数			3			



ときは入力信号に  $x_1$  と  $x_2$ , 5 入力の場合は入力信号に  $x_1 \sim x_5$ , 教師信号は両方とも  $x_1$  を用いる。  
また, BPN の出力関数には, 逆正接関数を用いる。

表 2.1 に学習精度評価計算機シミュレーション条件を示す。ここで, 結合荷重としきい値の初期値は, シミュレーション 1 回ごとに表 2.1 に示す範囲内で, 擬似乱数を用いて決定される。  
また,  $\alpha_{min}$  は, 粗探索で平均 2 乗平均平方根誤差 (Root Mean Square Error, 以下 RMSE と称する) が, 最小となる勾配法項の学習強化係数である。

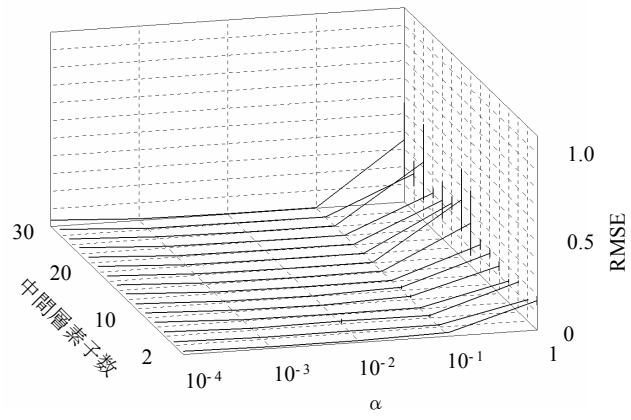
学習は, はじめに, 非線形信号予測システムとして用いられる BPN の中間層素子数を表 2.1 に示す範囲と間隔で設定し, それらに対して, 勾配法項の学習強化係数を表 2.1 に示す範囲と間隔で変化させながら, 3 回の粗探索を行い, その結果得られた平均 RMSE を比較する。  
次に, 最小平均 RMSE を実現した中間層素子数を用いて, 勾配法項の学習強化係数を表 2.1 に示す範囲と間隔で変化させながら, 3 回の細探索学習を行い, その結果得られた平均 RMSE を比較する。このようにして得られる両システムの学習結果を用いて, それらの学習精度の評価を行う。

## 2.2.2 結果

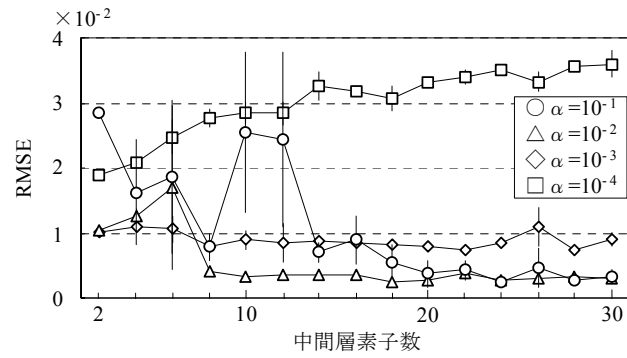
図 2.3~2.6 に, 中間層素子数と勾配法項の学習強化係数  $\alpha$  に対して学習で得られた平均 RMSE とその標準偏差を示す。ここで, 図中の縦棒は, 標準偏差の $\pm$ の範囲を示す。図 2.3 と 2.5 の(a)は, 粗探索学習後のすべての結果を 3 次元表示したものである。また, 同図(b)は, 同図(a) の平均 RMSE-中間層素子数面に対して,  $\alpha=1$  を除いて拡大表示したものである。図 2.3 より, 2 入力 BPN では, 中間層素子数 24,  $\alpha=0.1$  のとき最小平均 RMSE  $1.76 \times 10^{-3}$ , 図 2.5 より, 5 入力 BPN では, 中間層素子数 22,  $\alpha=0.1$  のとき最小平均 RMSE  $1.09 \times 10^{-3}$  が示されている。

図 2.4 と 2.6 は, 細探索学習後得られた最小平均 RMSE の付近を拡大して表示したものである。図 2.4 より, 2 入力 BPN では,  $\alpha=0.1$  のとき最小平均 RMSE  $1.76 \times 10^{-3}$ , 図 2.6 より, 5 入力 BPN では,  $\alpha=0.1$  のとき最小平均 RMSE  $1.09 \times 10^{-3}$  が示されている。両ネットワークとも, 粗探索学習と細探索学習での最小平均 RMSE が同じになっているが, これらは, 偶然に生じた結果である。この結果より, 5 入力 BPN は, 2 入力 BPN より最小平均 RMSE が 38.1%減少し, 学習精度が高いことが示されている。

表 2.2 は, 両ネットワークが細探索学習後に最小平均 RMSE を得たときのシミュレーショ



(a)



(b)

図 2.3 2 入力 BP ネットワークの粗探索学習後の平均 RMSE と標準偏差

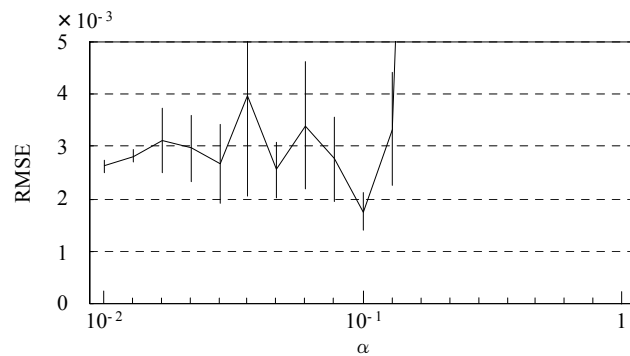
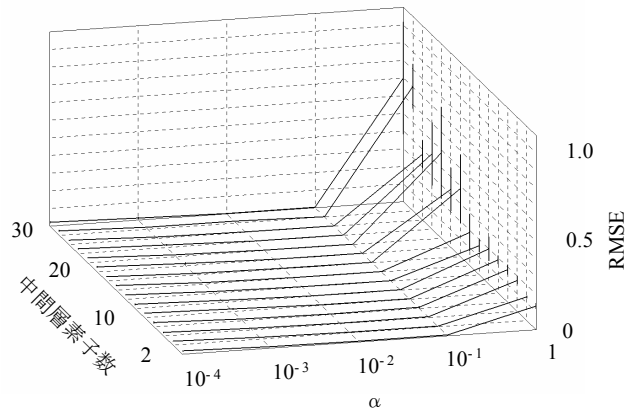
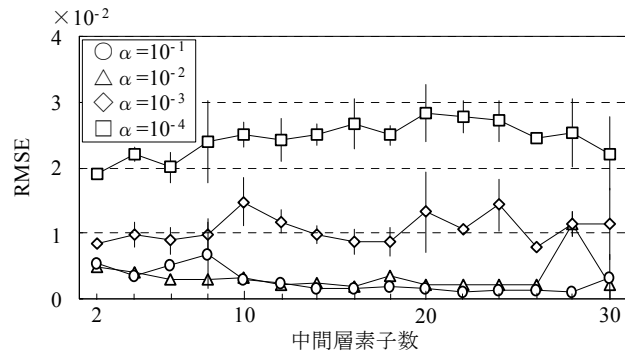


図 2.4 中間層素子数 24 の 2 入力 BP ネットワークの細探索学習後の平均 RMSE と標準偏差



(a)



(b)

図 2.5 5 入力 BP ネットワークの粗探索学習後の平均 RMSE と標準偏差

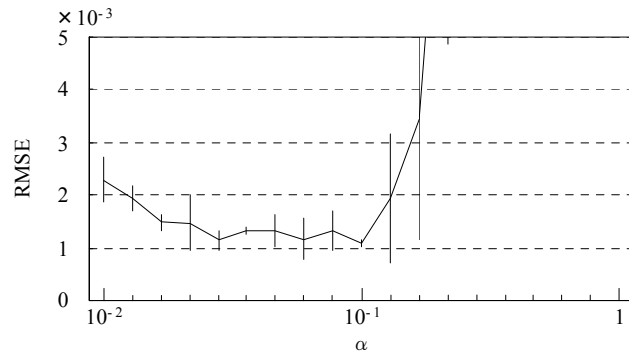


図 2.6 中間層素子数 22 の 5 入力 BP ネットワークの細探索学習後の平均 RMSE と標準偏差

表 2.2 細探索学習後の最小平均 RMSE

RMSE	入力数	2入力	5入力
	シミュレーション		
	1回目	$2.20 \times 10^{-3}$	$1.03 \times 10^{-3}$
	2回目	$1.32 \times 10^{-3}$	$1.19 \times 10^{-3}$
	3回目	$1.77 \times 10^{-3}$	$1.05 \times 10^{-3}$
平均		$1.76 \times 10^{-3}$	$1.09 \times 10^{-3}$
分散		$1.29 \times 10^{-7}$	$4.82 \times 10^{-9}$
標準偏差		$3.59 \times 10^{-4}$	$6.94 \times 10^{-5}$

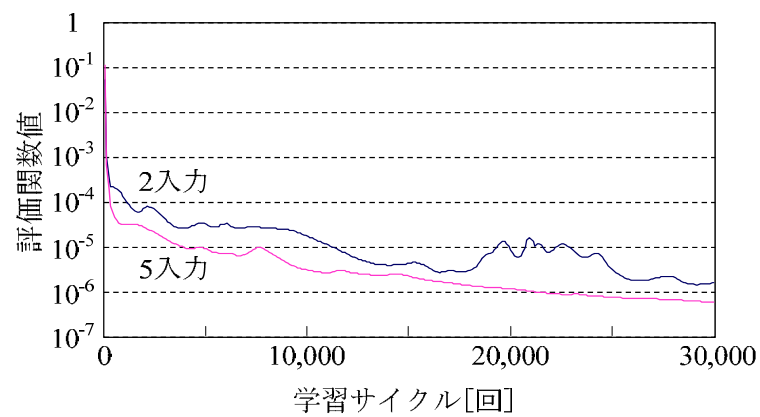
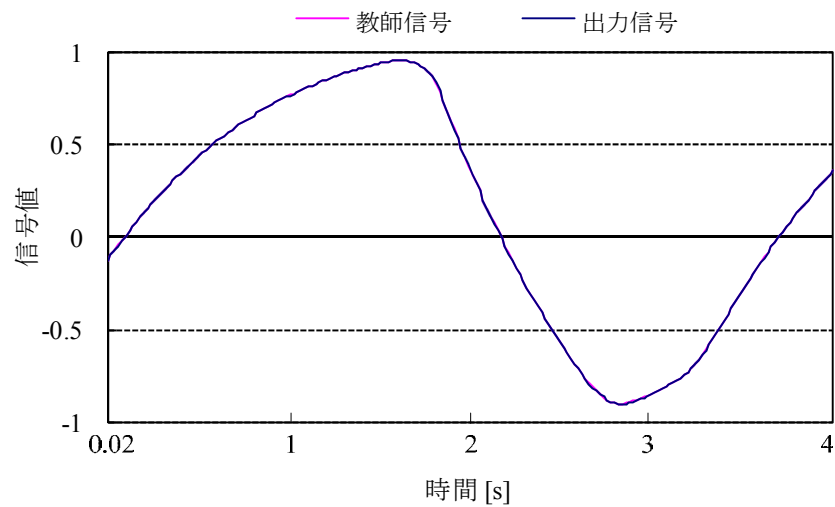
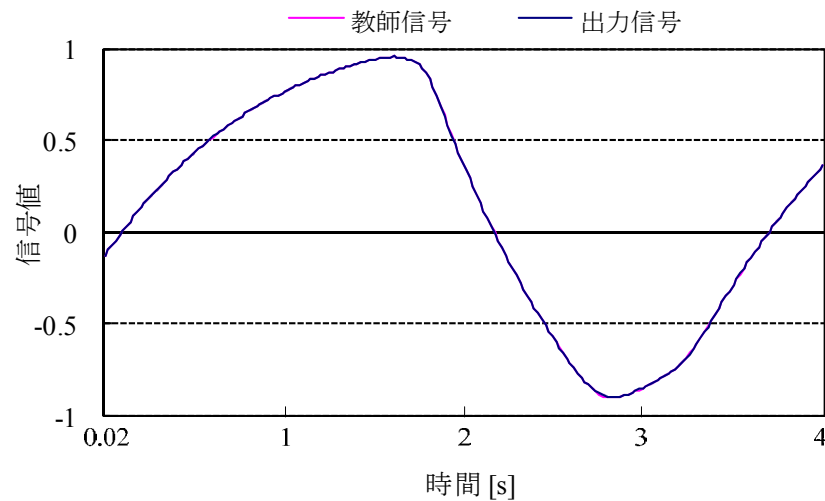


図 2.7 最小平均 RMSE のときの学習サイクルと 1 単位時間データ当りの平均評価関数値の平均との関係



(a) 2 入力 BP ネットワーク



(b) 5 入力 BP ネットワーク

図 2.8 細探索学習後の最小 RMSE を得たときのネットワークの出力信号波形

ン結果を示す．この表からも，5 入力 BPN は，2 入力 BPN より，最小 RMSE について 22.0% 減少し，学習精度が高いことが示されている．図 2.7 は，2 入力 BPN と 5 入力 BPN の細探索学習で，それぞれの最小平均 RMSE が得られたときの学習開始前，学習サイクル 1 回目，学習サイクル 100 回ごとに記録した学習サイクルと 1 単位時間データ当りの平均評価関数値の平均との関係を示す．この図より，5 入力 BPN の方が，2 入力 BPN より学習の起伏が少なく，安定した学習が得られていることが確認できる．図 2.8 は，最小平均 RMSE が得られた細探索学習の中で，最小 RMSE を実現した両ネットワークの出力信号波形と教師信号波形を比較したものを示す．両ネットワークとも視覚的には区別し難く，同等な出力信号が得られているように見える．

## 2.3 考察

はじめに，BPN の入力信号数による出力信号の実現能力の観点から検討する．教師信号と相関のある入力信号を付加することで，出力信号を生成するための基本となる信号成分が増える．すなわち，出力信号を生成するために都合がよい中間層の出力信号を生成しやすくなる．このため，BPN は，出力信号を実現する能力が向上することとなる．したがって，出力信号を生成するための基本となる信号成分が多い 5 入力 BPN の方が 2 入力 BPN より，学習精度が高いと

考えられる.

次に, BPN の入力信号数に対する入出力写像を学習する能力への影響の観点から検討する. BPN が, 与えられた入力信号値で, 1 つの出力信号値を決定できない場合, 1 つの出力信号値を決定できるように, さらに, その出力信号値を生成することができる適切な相関成分としての入力信号を付加しなければならない. BPN は, このように適切な入力信号を付加することで, 2 つ以上の出力信号値に対応する 1 つの入力信号値を分離する能力が高くなる. そして, BPN は, 1 つの入力信号値に対して 1 つの教師信号値を入出力写像として学習できる状態になる. 今回の計算機シミュレーションでは, 出力信号に対する入力信号の分離能力が向上する効果が, 2 入力 BPN の方より 5 入力 BPN の方に現れ, 5 入力 BPN の方が 2 入力 BPN より, 学習精度が高くなったと考えられる.

以上より, BPN へ教師信号と相関のある入力信号を付加することで, その学習精度が向上する場合があることが示された. この結果から, BPN で構成されたシステムやコントローラなどの学習能力に対する入力信号の不足がないかを再確認し, その不足があれば, さらにそれらが高性能となるように, 教師信号と相関のある適切な入力信号を付加して改善することが望まれる.

## 2.4 結言

2 入力 BPN と 5 入力 BPN を用いた 2 つの非線形信号予測システムの学習精度を評価した結果, 5 入力 BPN を用いた非線形信号予測システムの方が 2 入力 BPN を用いた非線形信号予測システムより, 最小平均 RMSE で 38.1%, 最小 RMSE で 22.0% 減少し, 学習精度が向上することが示された.

今後の課題は, BPN で構成されたシステムやコントローラなどの学習能力に対する入力信号の不足がないかを再確認し, その不足があれば, さらにそれらが高性能となるように, 教師信号と相関のある適切な入力信号を付加して改善することである.

## 参考文献

- [1] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning Representations by Back-propagating Errors," *Nature*, vol.323, no.6088, pp.533-536, 1986.
- [2] K. J. Lang and G. E. Hinton, "A Time-delay Neural Network Architecture for Speech Recognition," *CMU-CS-88-152*, pp.1-37, 1988.
- [3] J. S. Taur and S. Y. Kung, "Prediction-based Networks with ECG Application," *IEEE International Conference on Neural Networks*, vol.3, pp.1920-1925, 1993.
- [4] 小早川俊祐, 藤井崇文, 横井博一, "心電図に対するニューロンネットワークの非線形予測能力の評価," *BMFSA 第20回年次大会講演論文集*, pp.9-12, 2007.
- [5] 小早川俊祐, 藤井崇文, 横井博一, "心電図に対するボルテラニューロンネットワークの非線形予測能力の評価," *第9回日本知能情報ファジィ学会九州支部学術講演会予稿集*, pp.19-22, 2007.
- [6] 星文伸, 松坂知行, 木村昭穂, "FIR/IIR ニューラルネットワークによる時系列信号の予測に関する研究," *八戸工業大学紀要*, vol.23, pp.113-121, 2004.
- [7] 平雄一郎, 森元映治, 中村誠, "ニューラルネットを用いた魚種の識別--特徴量解析による入力数の低減," *水産大学校研究報告*, Vol.53, No.4, pp.137-145, 2005.

## 第3章 並列化によるニューロンネットワークの学習精度向上

### 3.1 緒言

プラントの入出力関係を明確な数学モデルとして表記できる場合は、BPN を用いやすい。しかし、それが未知の場合は、それを用い難い。例えば、物理的・数学的に関連性のない入出力間の写像を人工的に決定するような場合、すなわち、人間の脳の発想によるもの等は、このネットワークを用い難いと考えられる。この問題を解決する方法の一つとして、1 出力のユニットを並列に配置し、2 出力以上の BPN と同等なニューロンネットワークを新たに構成する PNN[1]-[5]がある。これにより、出力に対する有効な学習が、入力層を除く全ての層で行われ、前記問題を解決できるという経験が存在する。しかし、その理由と能力の評価は、いまだ定説化されておらず、経験的に良好であるという認識に過ぎない。ネットワークの並列型を適用した近年の応用研究には、M. Swiercz [6]らによる学習ベクトル量子化ニューロンネットワークを用いた中大脳動脈性発作の検出や I. Babaoglu [7]らによるニューロンネットワークを用いた冠動脈疾患の決定と損傷部位の予測などがあり、盛んに用いられるようになってきている。

本章では、PNN と BPN を用いた非線形時系列信号予測システムで、学習精度を比較し、PNN の優位性を示しながら、その理由を考察する。そして、誤差を収束させるための複出力関数システムの構成手法として、その結果を第 6 章で用いる。

### 3.2 並列型ニューロンネットワーク

#### 3.2.1 入出力特性

図 3.1 に離散時間の PNN を示す。PNN へ用いるニューロンは、様々なタイプを適用できる。本章では、PNN に通常の人工ニューロンを用いる。PNN の入出力特性を表すと、入力層では式(3.1)、中間層から出力層までは式(3.2)～(3.4)のように示される。

$$\begin{aligned} z_i \left\langle \begin{matrix} 1 \\ k \end{matrix} \right\rangle^{(\tau)} &= x_i \left\langle \begin{matrix} 1 \\ k \end{matrix} \right\rangle^{(\tau)} = x_k^{(\tau)} \\ (i &= 1, 2, \dots, n_{\text{out}}; k = 1, 2, \dots, n_{\text{in}}) \end{aligned} \quad (3.1)$$



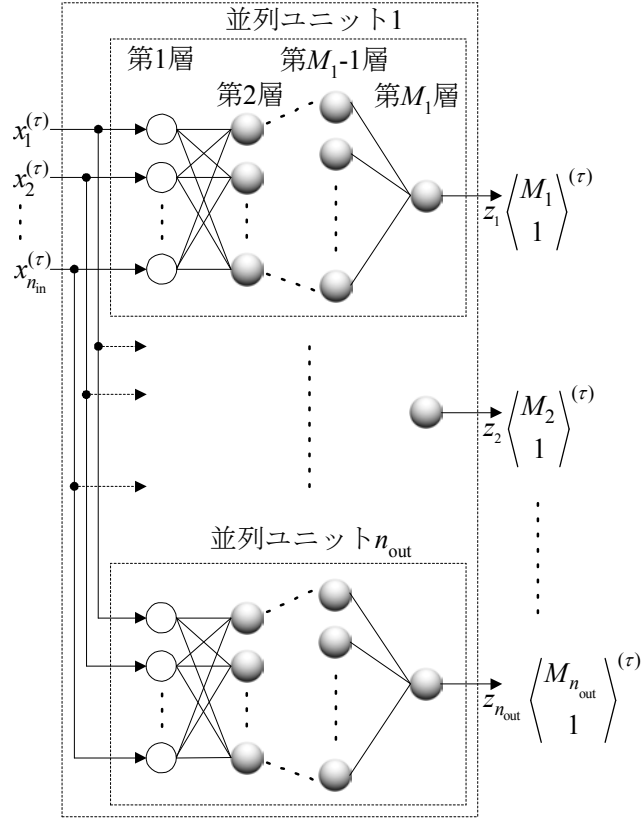


図 3.1 離散時間の並列型ニューロンネットワーク

$$u_i \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau)} = \sum_{j=1}^{n_{L-1}} w_i \left\langle \begin{matrix} L-1 & L \\ j & k \end{matrix} \right\rangle^{(\tau)} z_i \left\langle \begin{matrix} L-1 \\ j \end{matrix} \right\rangle^{(\tau)} \quad \left( \begin{matrix} L=2,3,\dots,M_i \\ i=1,2,\dots,n_{\text{out}}; j=1,2,\dots,n_{L-1}; k=1,2,\dots,n_L \end{matrix} \right) \quad (3.2)$$

$$s_i \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau)} = u_i \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau)} - h_i \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau)} \quad (L=2,3,\dots,M_i; i=1,2,\dots,n_{\text{out}}; k=1,2,\dots,n_L) \quad (3.3)$$

$$z_i \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau)} = f_i \left( s_i \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau)} \right) = A_i \tan^{-1} \left( s_i \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau)} \right) \quad (L=2,3,\dots,M_i; i=1,2,\dots,n_{\text{out}}; k=1,2,\dots,n_L) \quad (3.4)$$

ここで、括弧 $\langle \rangle$ 内は、上段が層番号、下段が素子番号、2 列表記のものは、左列が出力側素子、右列が入力側素子、 $x$ は入力信号、 $z$ は出力信号、 $w$ は結合荷重、 $u$ は入力荷重和、 $h$ はしきい値、 $s$ は入力の総和、 $f$ は出力関数、 $A$ は出力係数、 $M$ は出力層、各記号の添え字 $i$ は

並列ユニット番号である．また， $w$ と $h$ は，学習により変更される．

### 3.2.2 学習則

図 3.1 に示す PNN の学習則は，式(3.5)～(3.11)で示される．この学習則は，PNN の並列ユニットへ1出力の離散時間のBPネットワークへ適用される誤差逆伝搬則を適用したものである．

$$E_i = \frac{1}{2} \left( y_i - z_i \left\langle \begin{matrix} M_i \\ 1 \end{matrix} \right\rangle \right)^2 \quad (i=1,2,\dots,n_M) \quad (3.5)$$

$$\begin{aligned} \Delta w_i \left\langle \begin{matrix} L-1 & L \\ j & k \end{matrix} \right\rangle^{(\tau)} &= -\alpha_i \frac{\partial E_i^{(\tau)}}{\partial w_i \left\langle \begin{matrix} L-1 & L \\ j & k \end{matrix} \right\rangle^{(\tau)}} + \beta_i \Delta w_i \left\langle \begin{matrix} L-1 & L \\ j & k \end{matrix} \right\rangle^{(\tau-1)} \\ &= -\alpha_i r_i \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau)} z_i \left\langle \begin{matrix} L-1 \\ j \end{matrix} \right\rangle^{(\tau)} + \beta_i \Delta w_i \left\langle \begin{matrix} L-1 & L \\ j & k \end{matrix} \right\rangle^{(\tau-1)} \end{aligned}$$

$$\left( \begin{array}{l} L=2,3,\dots,M_i \\ i=1,2,\dots,n_M; j=1,2,\dots,n_{L-1}; k=1,2,\dots,n_L \end{array} \right) \quad (3.6)$$

$$w_i \left\langle \begin{matrix} L-1 & L \\ j & k \end{matrix} \right\rangle^{(\tau+1)} = w_i \left\langle \begin{matrix} L-1 & L \\ j & k \end{matrix} \right\rangle^{(\tau)} + \Delta w_i \left\langle \begin{matrix} L-1 & L \\ j & k \end{matrix} \right\rangle^{(\tau)}$$

$$\left( \begin{array}{l} L=2,3,\dots,M_i \\ i=1,2,\dots,n_M; j=1,2,\dots,n_{L-1}; k=1,2,\dots,n_L \end{array} \right) \quad (3.7)$$

$$\begin{aligned} \Delta h_i \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau)} &= -\alpha_i \frac{\partial E_i^{(\tau)}}{\partial h_i \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau)}} + \beta_i \Delta h_i \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau-1)} \\ &= \alpha_i r_i \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau)} + \beta_i \Delta h_i \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau-1)} \end{aligned}$$

$$(L=2,3,\dots,M_i; i=1,2,\dots,n_M; k=1,2,\dots,n_L) \quad (3.8)$$

$$h_i \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau+1)} = h_i \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau)} + \Delta h_i \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau)} \quad (L=2,3,\dots,M_i; i=1,2,\dots,n_M; k=1,2,\dots,n_L) \quad (3.9)$$

$$\begin{aligned}
r_i \left\langle \begin{matrix} M_i \\ 1 \end{matrix} \right\rangle^{(\tau)} &= \frac{\partial E_i^{(\tau)}}{\partial s_i \left\langle \begin{matrix} M_i \\ 1 \end{matrix} \right\rangle^{(\tau)}} \\
&= - \left( y_i^{(\tau)} - z_i \left\langle \begin{matrix} M_i \\ 1 \end{matrix} \right\rangle^{(\tau)} \right) \times A_i \frac{1}{1 + \left[ s_i \left\langle \begin{matrix} M_i \\ 1 \end{matrix} \right\rangle^{(\tau)} \right]^2} \\
(i=1, 2, \dots, n_M) &
\end{aligned} \tag{3.10}$$

$$\begin{aligned}
r_i \left\langle \begin{matrix} L \\ j \end{matrix} \right\rangle^{(\tau)} &= \frac{\partial E_i^{(\tau)}}{\partial s_i \left\langle \begin{matrix} L \\ j \end{matrix} \right\rangle^{(\tau)}} \\
&= A_i \frac{1}{1 + \left[ s_i \left\langle \begin{matrix} L \\ j \end{matrix} \right\rangle^{(\tau)} \right]^2} \times \sum_{k=1}^{n_{L+1}} \left( r_i \left\langle \begin{matrix} L+1 \\ k \end{matrix} \right\rangle^{(\tau)} w_i \left\langle \begin{matrix} L & L+1 \\ j & k \end{matrix} \right\rangle^{(\tau)} \right) \\
\left( \begin{matrix} L=2, 3, \dots, M_i-1 \\ i=1, 2, \dots, n_M; j=1, 2, \dots, n_L; k=1, 2, \dots, n_{L+1} \end{matrix} \right) &
\end{aligned} \tag{3.11}$$

ここで、 $E$  は評価関数、 $y$  は教師信号、 $\Delta w$  は結合荷重の変更量、 $\Delta h$  はしきい値の変更量、 $\alpha$  は勾配法項の強化係数、 $\beta$  は慣性項の強化係数、 $r$  は強化信号である。

### 3.3 計算機シミュレーション

#### 3.3.1 方法

図 3.2 に、時刻  $\tau$  の入力に対して、時刻  $\tau+1$  の予測を得る 5 入力 4 出力の非線形時系列信号予測システムを示す。このシステムへ各 3 層の BPN と PNN を適用し、表 3.1 の条件下で、最小平均 RMSE を求め、その値を比較する計算機シミュレーションを行う。入力信号と教師信号は、第 2 章の図 2.2 の学習用非線形プラント信号を用いる。結合荷重としきい値の初期値は、表 3.1 に示す範囲内で、学習計算機シミュレーション 1 回ごとに、疑似乱数を用いて決定される。中間層素子数の範囲は、BPN と PNN の 1 出力当りの素子数が同じになるように定める。また、 $\alpha_{min1}$  と  $\alpha_{min2}$  は、それぞれ粗探索学習と細探索学習で平均 RMSE が最小となる勾配法項の学習強化係数である。

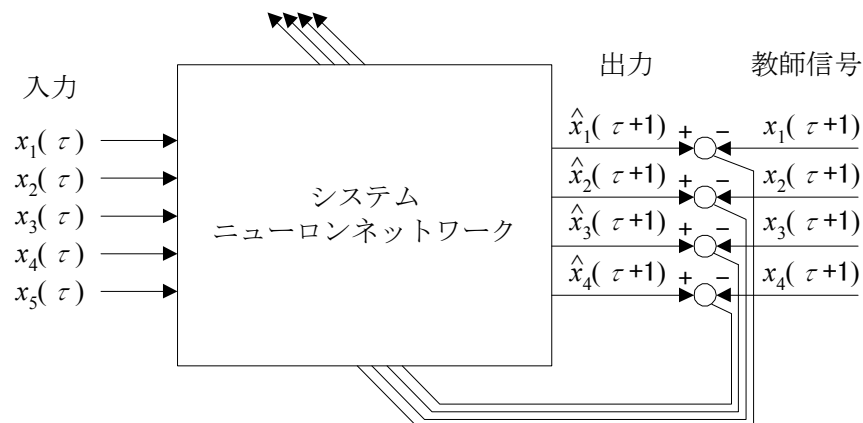


図 3.2 5 入力 4 出力非線形時系列信号予測システム

表 3.1 学習精度評価計算機シミュレーション条件

種類			BPN		PNN (1ユニット)	
条件項目						
学習則			誤差逆伝搬則		PNNの学習則	
初期値		結合荷重		-0.3～0.3		
		しきい値		-0.3～0.3		
中間層素子数			範囲		2～216	
			間隔		2～54	
			2			
学習強化係数	勾配法項	粗探索	範囲		$10^{-4} \sim 1$	
			間隔		10倍	
		細探索	範囲		$0.2 \alpha_{min1} \sim 0.9 \alpha_{min1}, 2 \alpha_{min1} \sim 9 \alpha_{min1}$	
			間隔		$0.1 \alpha_{min1}, \alpha_{min1}$	
		細々探索	範囲		$\alpha_{min2}$ を中心に細探索で使した前後の間隔	
			間隔		上記間隔の10分の1	
慣性項			0			
学習サイクル			30,000			
施行回数			4			

学習は、はじめに、非線形時系列信号予測システムの間層素子数を表 3.1 に示す範囲と間隔で設定し、勾配法項の学習強化係数を表 3.1 に示す範囲と間隔で変化させながら、4 回の粗探索学習を行う。その結果得られた平均 RMSE を比較する。次に、粗探索学習で最小平均 RMSE を実現した中間層素子数と勾配法項の学習強化係数を用いて、勾配法項の学習強化係数を表 3.1 に示す範囲と間隔で変化させながら、4 回の細探索学習を行う。さらに、細探索学習で最小平均 RMSE を実現した勾配法項の学習強化係数を用いて、勾配法項の学習強化係数を表 3.1 に示す範囲と間隔で変化させながら、4 回の細々探索学習を行う。このようにして得られる両システムニューロンネットワークの最小平均 RMSE を用いて、学習精度の評価を行う。

### 3.3.2 結果

図 3.3～3.8 に、BPN と PNN の粗探索、細探索、細々探索の学習後に得られた平均 RMSE とその標準偏差を示す。ここで、図中の縦棒は、標準偏差の $\pm$ の範囲を示す。表 3.2 に BPN と PNN の最小平均 RMSE 時の条件を示す。この表の各波形について、最小平均 RMSE を比較した結果、PNN のすべての学習精度は、BPN より高いことが確認できる。また、PNN のすべての出力の最小平均 RMSE の平均値は、BPN より 39.0%減少していることが示されている。

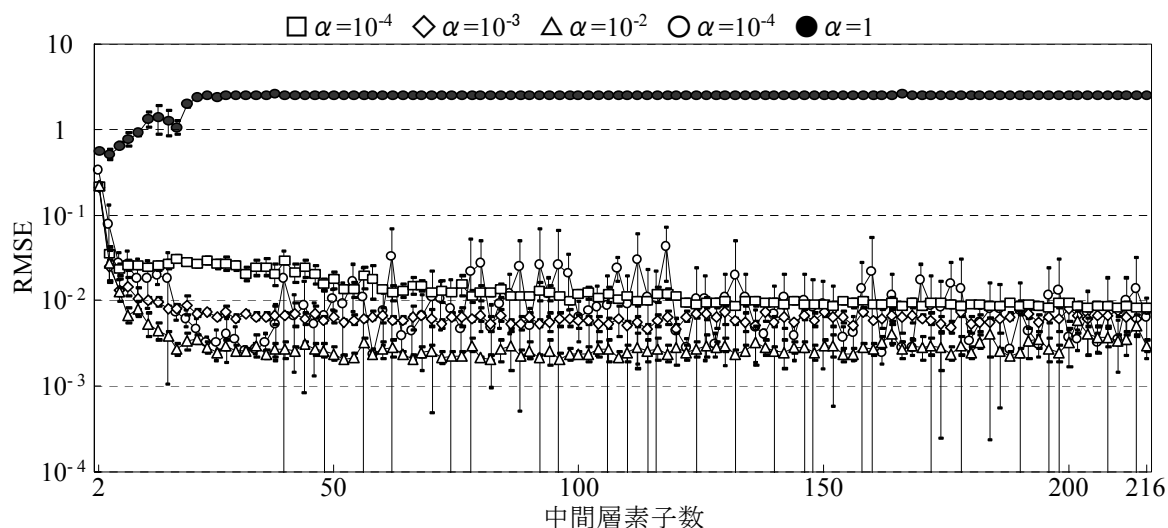
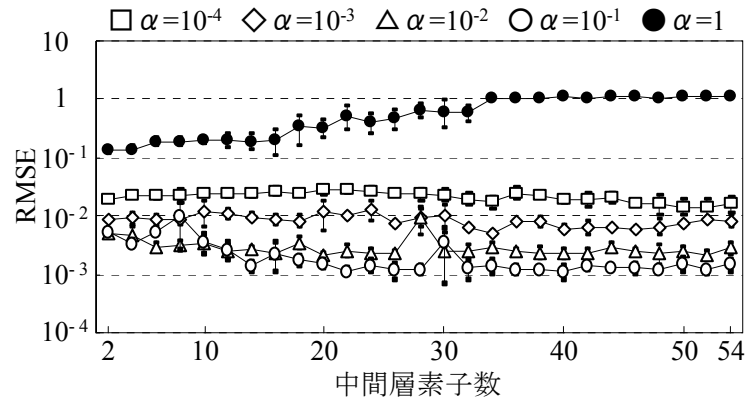
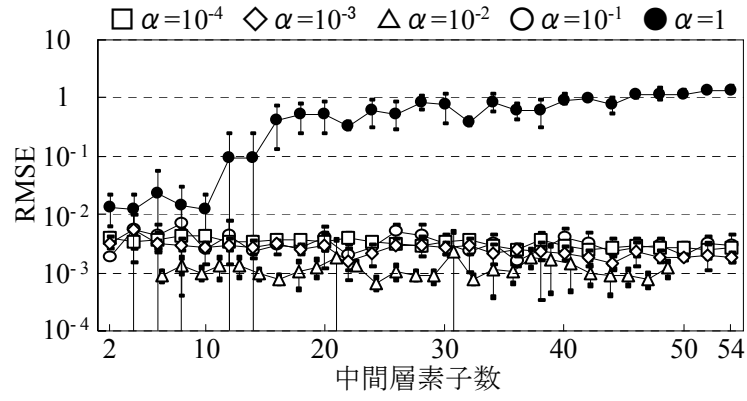


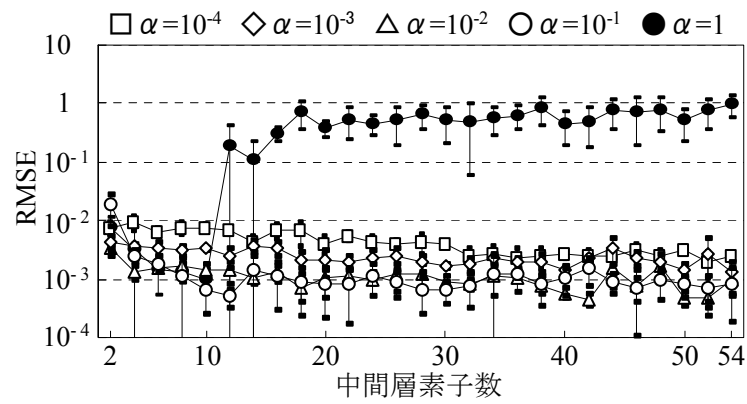
図 3.3 BP ネットワークの粗探索学習後の平均 RMSE と標準偏差



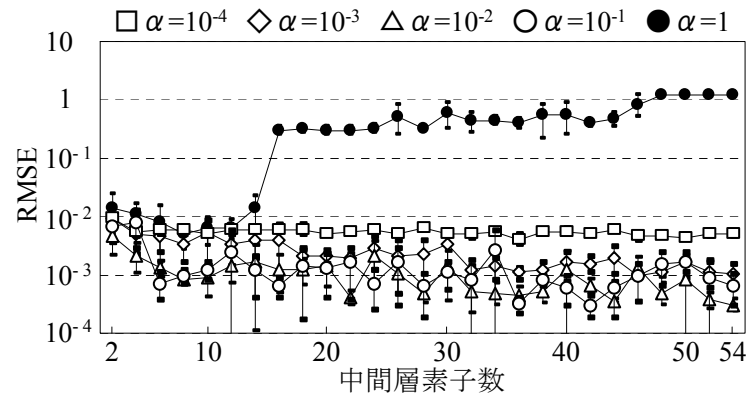
(a)  $x_1$



(b)  $x_2$



(c)  $x_3$



(d)  $x_4$

図 3.4 並列型ニューロンネットワークの粗探索学習後の平均 RMSE と標準偏差

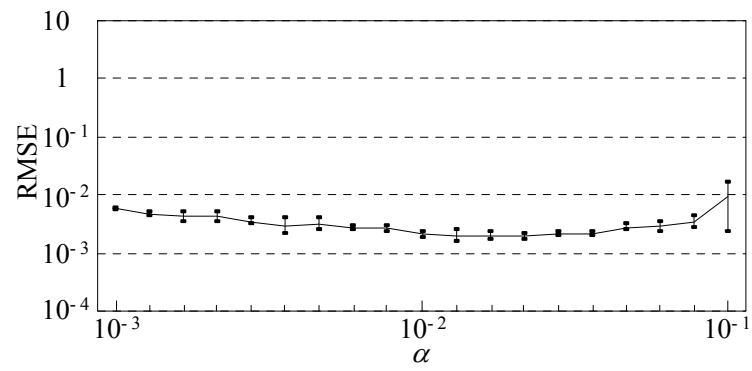
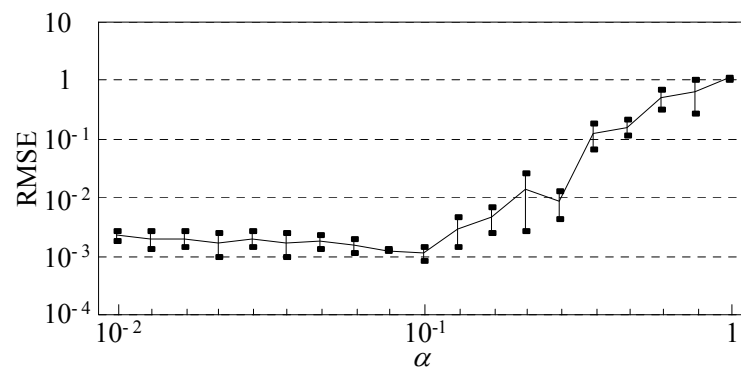
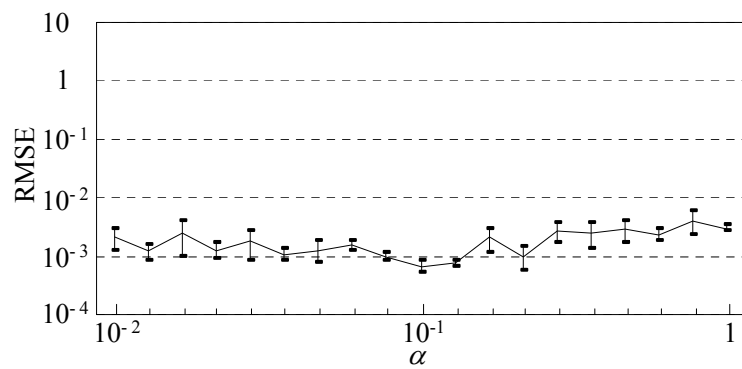


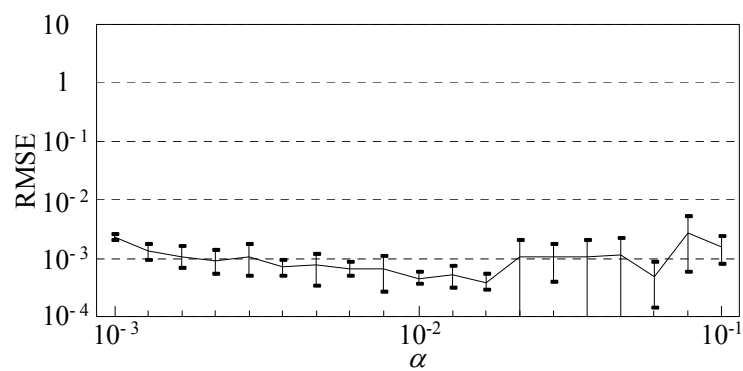
図 3.5 BP ネットワークの細探索学習後の平均 RMSE と標準偏差



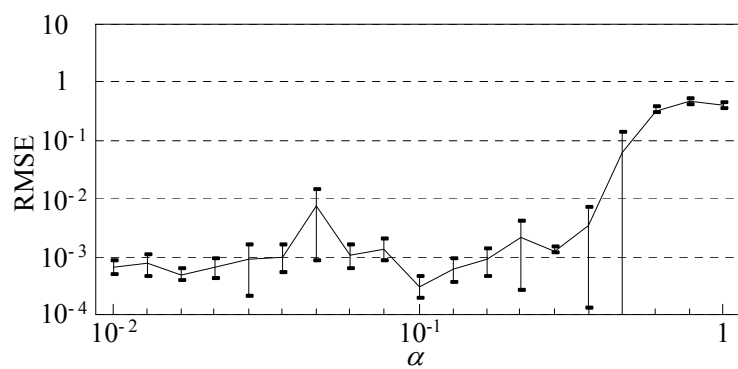
(a)  $x_1$



(b)  $x_2$



(c)  $x_3$



(d)  $x_4$

図 3.6 並列型ニューロンネットワークの細探索学習後の平均 RMSE と標準偏差



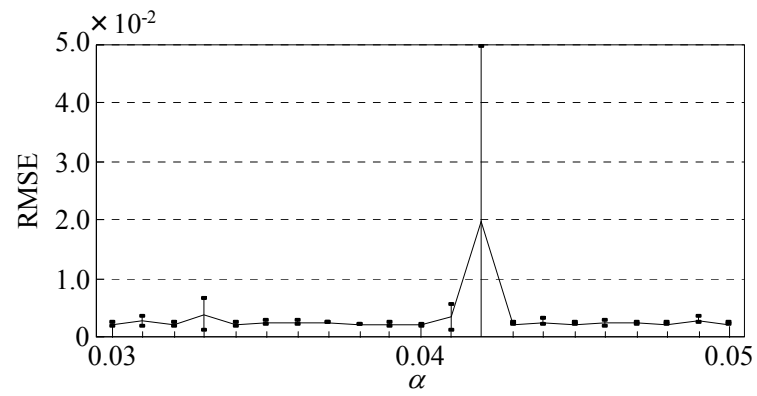
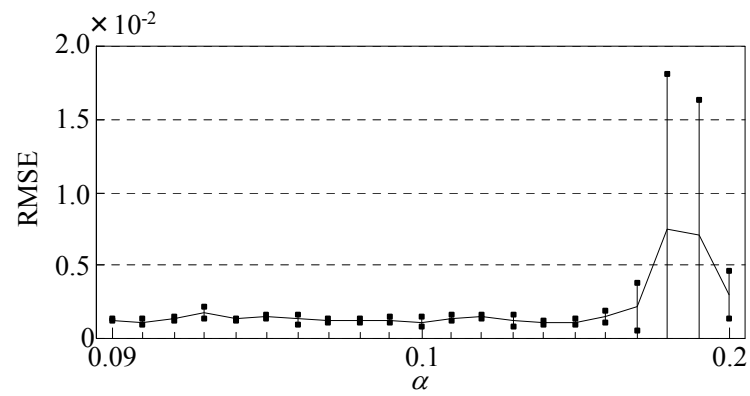
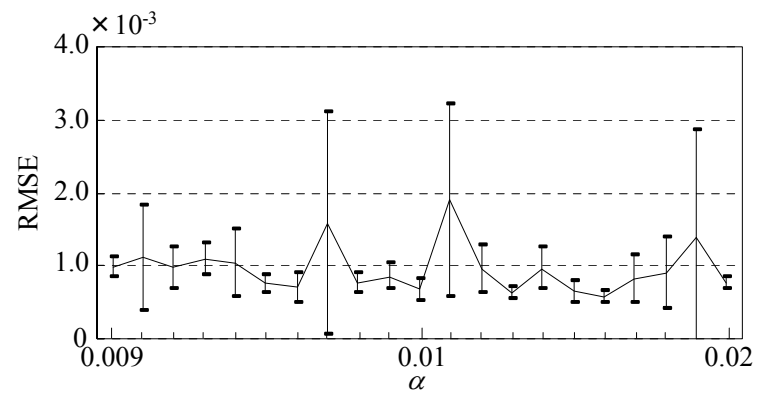


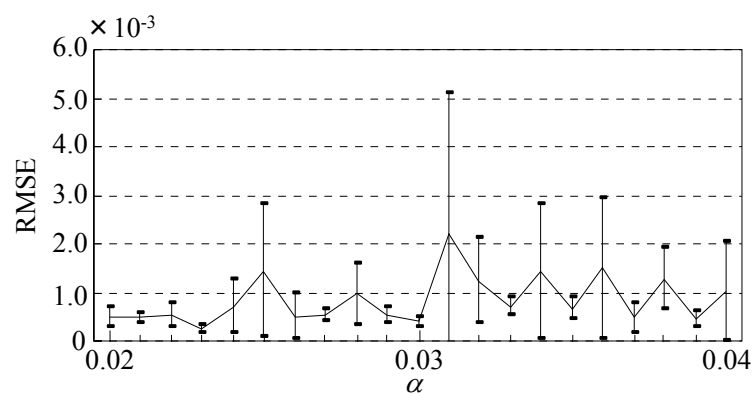
図 3.7 BP ネットワークの細々探索学習後の平均 RMSE と標準偏差



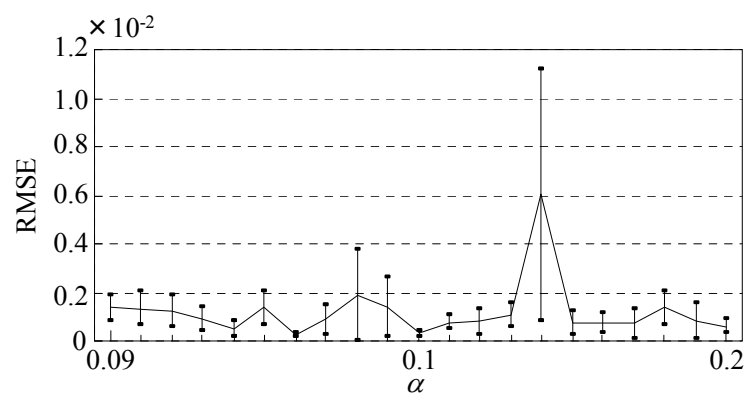
(a)  $x_1$



(b)  $x_2$



(c)  $x_3$



(d)  $x_4$

図 3.8 並列型ニューロンネットワークの細々探索学習後の平均 RMSE と標準偏差

表 3.2 細々探索学習後の最小平均 RMSE 時の条件

(a) BP ネットワーク

データ名 \ 波形名	$x_1$	$x_2$	$x_3$	$x_4$
中間層素子数	52			
1出力当りの中間層素子数	13			
学習強化係数	$4.0 \times 10^{-2}$			

(b) 並列型ニューロンネットワーク

データ名 \ 波形名	$x_1$	$x_2$	$x_3$	$x_4$
中間層素子数	40	24	42	42
合計中間層素子数	148			
1出力当りの中間層素子数	37			
学習強化係数	$1.4 \times 10^{-1}$	$1.6 \times 10^{-2}$	$2.3 \times 10^{-2}$	$9.6 \times 10^{-2}$

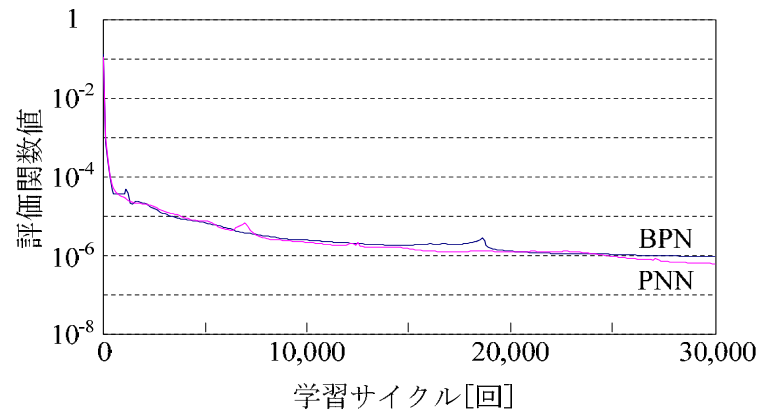
表 3.3 細々探索学習後の最小平均 RMSE

(a) BP ネットワーク

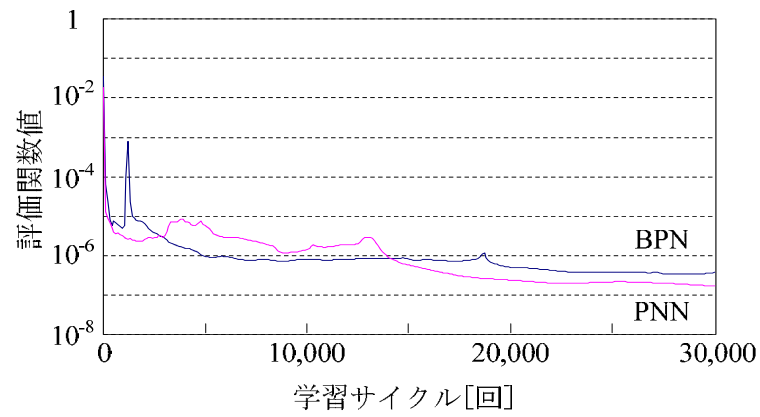
RMSE \ 波形名		$x_1$	$x_2$	$x_3$	$x_4$	平均
シミュレーション	1回目	$1.56 \times 10^{-3}$	$7.65 \times 10^{-4}$	$7.46 \times 10^{-4}$	$6.11 \times 10^{-4}$	$9.20 \times 10^{-4}$
	2回目	$1.36 \times 10^{-3}$	$8.93 \times 10^{-4}$	$6.00 \times 10^{-4}$	$7.82 \times 10^{-4}$	$9.10 \times 10^{-4}$
	3回目	$1.29 \times 10^{-3}$	$7.81 \times 10^{-4}$	$7.54 \times 10^{-4}$	$3.48 \times 10^{-4}$	$7.92 \times 10^{-4}$
	4回目	$1.30 \times 10^{-3}$	$9.89 \times 10^{-4}$	$8.71 \times 10^{-4}$	$1.04 \times 10^{-3}$	$1.05 \times 10^{-3}$
平均		$1.38 \times 10^{-3}$	$8.57 \times 10^{-4}$	$7.43 \times 10^{-4}$	$6.95 \times 10^{-4}$	$9.18 \times 10^{-4}$
分散		$1.18 \times 10^{-8}$	$8.23 \times 10^{-9}$	$9.26 \times 10^{-9}$	$6.32 \times 10^{-8}$	$5.28 \times 10^{-9}$
標準偏差		$1.09 \times 10^{-4}$	$9.07 \times 10^{-5}$	$9.62 \times 10^{-5}$	$2.51 \times 10^{-4}$	$7.26 \times 10^{-5}$

(b) 並列型ニューロンネットワーク

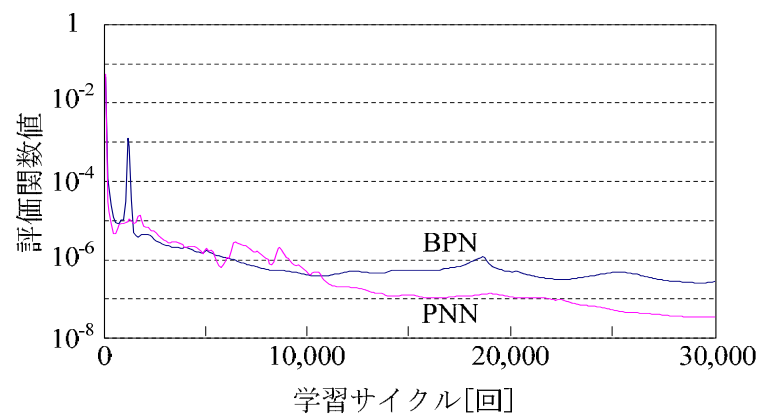
RMSE \ 波形名		$x_1$	$x_2$	$x_3$	$x_4$	平均
シミュレーション	1回目	$8.84 \times 10^{-4}$	$6.68 \times 10^{-4}$	$2.50 \times 10^{-4}$	$1.46 \times 10^{-4}$	$4.87 \times 10^{-4}$
	2回目	$1.29 \times 10^{-3}$	$4.80 \times 10^{-4}$	$1.56 \times 10^{-4}$	$2.09 \times 10^{-4}$	$5.34 \times 10^{-4}$
	3回目	$9.99 \times 10^{-4}$	$5.15 \times 10^{-4}$	$1.78 \times 10^{-4}$	$3.29 \times 10^{-4}$	$5.05 \times 10^{-4}$
	4回目	$1.27 \times 10^{-3}$	$6.36 \times 10^{-4}$	$3.90 \times 10^{-4}$	$3.58 \times 10^{-4}$	$6.64 \times 10^{-4}$
平均		$1.11 \times 10^{-3}$	$5.75 \times 10^{-4}$	$2.44 \times 10^{-4}$	$2.61 \times 10^{-4}$	$5.48 \times 10^{-4}$
分散		$3.02 \times 10^{-8}$	$6.27 \times 10^{-9}$	$8.36 \times 10^{-9}$	$7.47 \times 10^{-9}$	$1.88 \times 10^{-9}$
標準偏差		$1.74 \times 10^{-4}$	$7.92 \times 10^{-5}$	$9.14 \times 10^{-5}$	$8.64 \times 10^{-5}$	$4.34 \times 10^{-5}$



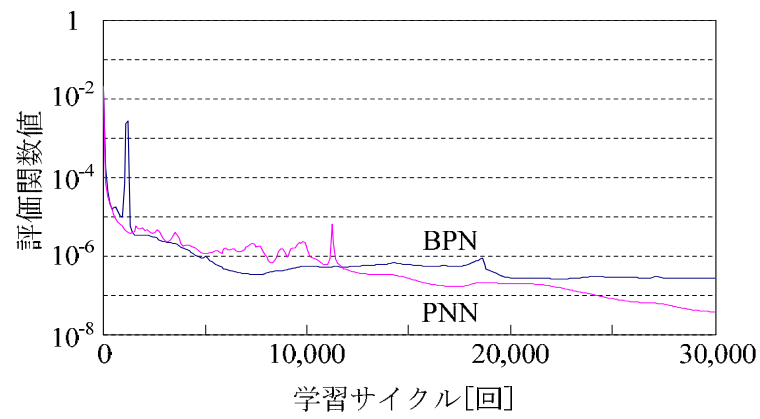
(a)  $x_1$



(b)  $x_2$



(c)  $x_3$



(d)  $x_4$

図 3.9 最小平均 RMSE のときの学習サイクルと 1 単位時間データ当りの平均評価関数値の平均の関係

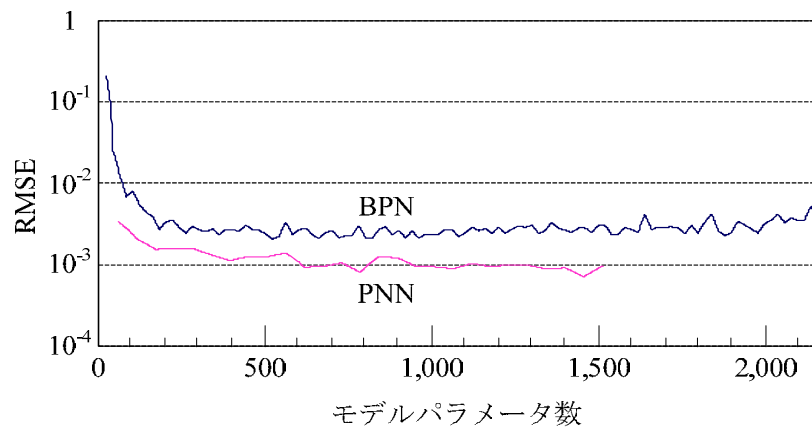
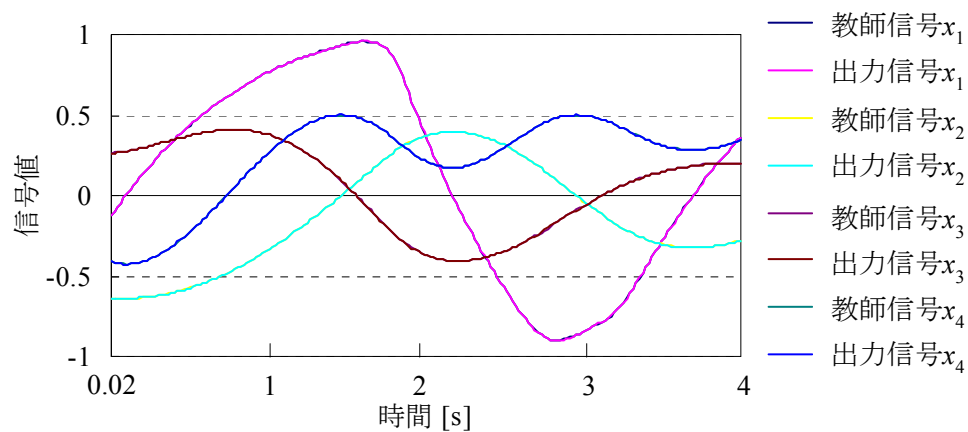
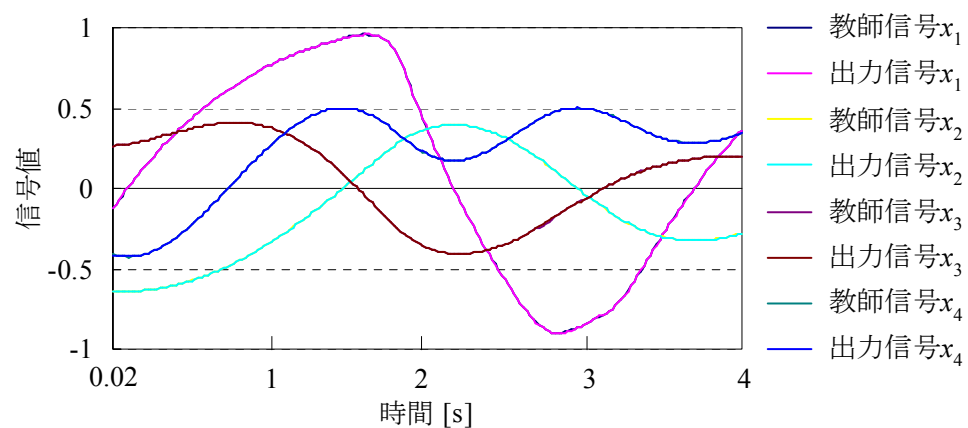


図 3.10 粗探索学習後の総モデルパラメータ数と 1 出力当りの最小平均 RMSE の関係



(a) BP ネットワーク



(c) 並列型ニューロンネットワーク

図 3.11 細々探索学習後の最小 RMSE を得たときのネットワークの出力信号波形

表 3.3 は、両ネットワークが細々探索学習後に最小平均 RMSE を得たときのシミュレーション結果を示す。この表からも、PNN は、BPN より、最小 RMSE について 38.5%減少し、学習精度が高いことが示されている。図 3.9 は、BPN と PNN の細々探索学習で、それぞれの最小平均 RMSE が得られたときの学習開始前、学習サイクル 1 回目、学習サイクル 100 回ごとに記録した学習サイクルと 1 単位時間データ当りの平均評価関数値の平均の関係を示す。この図より、学習サイクル 1 万回前後までは、BPN の方が、PNN より学習が良好であるが、それ以上の学習サイクルでは逆転し、PNN の学習精度の方が良好になっていることが確認できる。図 3.10 は、粗探索学習後の総モデルパラメータ数と 1 出力当りの最小平均 RMSE の関係を示す。

この図より、粗探索学習では、どの総モデルパラメータ数においても PNN の方が、BPN より学習精度が優れていることが確認できる。図 3.11 は、最小平均 RMSE が得られた細々探索学習の中で、最小 RMSE を実現した両ネットワークの出力信号波形と教師信号波形を比較したものを示す。両ネットワークとも視覚的には区別し難く、同等な出力信号が得られているように見える。

### 3.4 考察

はじめに、学習則の観点から考察する。2 出力以上の BPN の最終中間層では、式(3.6), (3.8) の勾配法項から導かれる強化信号  $r$  の計算式(3.11)の中で、結合荷重としきい値の変更量への干渉の原因となる出力層の全ての出力についての強化信号と結合荷重の積の和の計算が発生する。これは、任意の出力が学習を完全に収束したとしても、学習が収束していない他の出力によって、最終中間層の結合荷重やしきい値が、変更させられるということである。したがって、学習を完全に収束した出力が、この変更により、その状態を連続して保つということは、非常に難しいと考えられる。すなわち、この出力は、教師信号に対して振動的になる。さらに、この振動が、他の出力にも振動的な影響を与える。これが、出力間で生じる学習の相互干渉である。この干渉を含んだ強化信号を基に、入力層へ向かって、中間層内で次々と結合荷重やしきい値の変更計算がされるため、干渉も逆伝播することになる。PNN は、1 つの並列ユニットに対して 1 出力のため、前記の干渉がなく、良好なニューロンネットワークの能力が期待できるものになる。

次に、中間層の共有化の観点から考察する。BPN は、各出力に対応する中間層を共有している。出力ごとに用意されている中間層を 1 つにして、各出力で共有することができる条件、すなわち、中間層の共有条件は、すべての入力信号ベクトルに対して、出力ごとに用意されている中間層の出力信号ベクトルが等しくなる場合である。この中間層の共有条件は、出力ごとに用意されている中間層の素子の係数の構成が、どの中間層においても等しいならば、明らかに満たされる。しかし、実際には、このような条件を満たすことは、非常に困難である。このため、BPN は、中間層の共有条件を満たさない中間層の出力信号ベクトルを用いて、目的とする出力を得るための学習能力を出力層に十分持っていないなければならない。これでは、出力層だけを学習するパーセプトロンの学習と性質が似たままになってしまう。一方、PNN は、BPN が各出力に対応する中間層を共有し難いという前記の問題を回避するために、出力ごとに中間層を

用意している．このため，1 つの出力の学習のために，学習が中間層と出力層で行われ，良好なニューロンネットワークの能力が期待できるものになる．

最後に，総モデルパラメータ数の観点から考察する．一般に，モデルパラメータ数が多いほど，ニューロンネットワークの実現能力は高くなる．したがって，同じニューロンで構成されている異なるニューロンネットワークでも，同じ総モデルパラメータ数であれば，同じ実現能力があると考えられる．しかし，粗探索時の結果は，どの総モデルパラメータにおいても，PNNの方がBPNより高い学習精度を示している．これは，前記のBPNの出力間で生じる学習の相互干渉や中間層共有化の困難性により，BPNの学習能力が劣っているためと考えられる．

以上の観点から，PNNの学習精度は，BPNより常に高くなると考えられる．この結果により，BPNを用いている従来のアプリケーションを，PNNを用いたものにするだけで，学習精度を改善できる．また，PNNの各並列ユニットへ第6章のECNN[12],[13]を適用したECPNNにすることで，さらに学習精度を改善できると考えられる．

### 3.5 結言

BPNとPNNを5入力4出力非線形時系列信号予測システムへ適用し，非線形プラントから得られた非線形時系列信号を用いて，それらの学習精度を比較した．その結果，PNNは，BPNより，平均RMSEで39.0%，最小RMSEで38.5%減少していた．このとき，PNNの学習精度は，BPNより，すべての出力において高かった．この原因は，2出力以上のBPNにおいて，出力間で生じる学習の干渉や各出力で用いる中間層の共有化実現の困難性にあると解明した．今後の課題は，PNNをさらに改良したECPNN[8]の計算機シミュレーションを行い，その有効性を検討することである．



## 参考文献

- [1] 小早川俊祐, 横井博一, “航空機ピッチ姿勢制御のためのボルテラフィルタ内蔵型ニューラルネットワーク,” 平成17年度電気関係学会九州支部連合大会講演論文集, p.429, 2005.
- [2] 小早川俊祐, 横井博一, “並列化ニューロンネットワークの予測能力評価,” 第8回日本知能情報ファジィ学会九州支部学術講演会予稿集, pp.3-6, 2006.
- [3] 小早川俊祐, 横井博一, “並列化ニューロンネットワークの航空機における予測問題への適用,” 信学技報, vol.106, no.471, SANE2006-119-133, pp.43- 45, 2007.
- [4] S. Kobayakawa and H. Yokoi, “Evaluation of the Learning Capability of a Parallel-type Neuron Network,” Proceedings 1<sup>st</sup> International Symposium Information and Computer Elements 2007, pp.43-47, 2007.
- [5] 小早川俊祐, 横井博一, “並列型ニューロンネットワークの予測出力精度優位性の実験的検討,” 信学技報, vol.108, no. 54, NC2008-1-10, pp.29-34, 2008.
- [6] M. Swiercz, J. Kochanowicz, J. Weigele, R. Hurst, D. S. Liebeskind, Z. Mariak, E. R. Melhem, and J. Krejza, “Learning Vector Quantization Neural Networks Improve Accuracy of Transcranial Color-coded Duplex Sonography in Detection of Middle Cerebral Artery Spasm—Preliminary Report,” Neuroinformatics, vol.6, no.4, pp.279-290, 2008.
- [7] I. Babaoglu, O. K. Baykan, N. Aygul, K. Ozdemir, and M. Bayrak, “Assessment of Exercise Stress Testing with Artificial Neural Network in Determining Coronary Artery Disease and Predicting Lesion Localization,” Expert Systems with Applications, vol.36, issue 2, Part 1, pp.2562-2566, 2009.
- [8] S. Kobayakawa and H. Yokoi, “Proposal of Error Convergence-type Neuron Network System,” Presented proceeding to 2008 International Symposium on Intelligent Informatics, Kumamoto, Japan, pp.1-10, 2008.

## 第4章 ボルテラフィルタ内蔵によるニューロンネットワークの学習精度向上

### 4.1 緒言

心電図をはじめとするデジタル記録された生体信号のデータ量は、膨大であることが多く、そのデータを保管するためのデータ圧縮法が、デジタル時代の幕開けと共に要求されてきた。心電図のデータ圧縮は、折れ線近似を用いた手法に始まるが、これらの手法では、心電図に重畳して計測される節電図などの雑音に比較的弱く、再構成された心電図波形が、不自然であるなどの問題がある。そこで、情報理論の立場から多くの心電図データ圧縮法が、提案されてきた。

データ圧縮の基本は、データの冗長性を削減することである。心電図に関しては、近い時刻の標本点間や心拍間の類似性を利用した手法として、直交変換やエルミート関数[1]を用いた手法、直交ウェーブレットフィルタ[2]-[4]などを利用したサブバンド符号化[5]、それらを組み合わせた手法、DPCMなどの線形予測符号化を用いた手法などが提案されている。これらの手法は、安定した心拍が続く場合、高い圧縮性能が得られるが、発作的に出現する調律異常波形に対しては、圧縮性能の急激な低下や1心拍の切出しが困難になるなどの問題が生じる。この問題へ対処するために、長期予測を用いた手法[6]やBurrows-Wheeler変換と線形予測の逆ランクを用いた手法[7]が提案されている。また、最近の圧縮方法として、可逆圧縮ができる反辞書式予測符号化[8]が提案されている。

近い時刻の標本点間の類似性を利用した手法である変換符号化やサブバンド符号化は、一般に、変換係数、あるいは、サブバンドごとにエントロピー符号器を設計しなければならない。一方、DPCMをはじめとする線形予測符号化は、予測誤差系列に対する唯一のエントロピー符号器を用意するだけで済み、取り扱いやすいことを特徴とする。また、予測符号化は、変換符号化やサブバンド符号化と比較して、低ビットレート域での符号化性能に劣る。しかし、この方法は、高ビットレート域で、ほぼ同等の性能を有することから、主に高ビットレート域でのデータ圧縮に利用されている。

通常用、または、波形劣化が認められない程度の再構成精度が要求される研究用の心電図データベースは、高ビットレート域でのデータ圧縮が要求される。このため、これらの用途では、DPCMをはじめとする線形自己回帰モデルを予測器に用いた予測符号化の適用が提案され

てきた。しかし、心電図の予測符号化は、予測次数を2, 3次と上げていっても、予測誤差がほとんど減少せず、圧縮性能が改善しないことが指摘されている。さらに、心電図波形において、診断時に最も重要な心臓の収縮を反映するQRS波の予測誤差が、ほかの部分と比較して、特に大きくなるという問題がある。その理由としては、心電図、特にQRS波の生成について、非線形なメカニズムが作用し、線形予測モデルでは表現しきれない非線形な成分が、QRS波に含まれているからであると考えられている。したがって、心電図の予測には、非線形な予測モデルを含む非線形予測[9]-[19]が必要となる。その中に、心電図に対して非線形予測能力を有しているボルテラ汎関数級数や、それより少し汎化精度が高いIDNN [17][19]-[24]がある。しかし、これらを用いても、全体的な予測誤差の減少は見られるものの、QRS波に関しては十分ではない。

ところで、VNNは、非線形性の強い時系列信号の処理能力の高いNNとして、横井によって考案され、ヒューマノイドロボット[25]-[27]、筋電義手[28]、自動車の横滑り[29]、航空機の縦系姿勢[30]の各種制御、脳波アーチファクトの除去[31]などへ応用した研究で、その有効性が確認されている。3層全てが2次ボルテラニューロン（Second-order Volterra Neuron, 以下2VNと称する）から構成される2VNNは、1次ボルテラニューロン（First-order Volterra Neuron, 以下1VNと称する）から構成される1VNNより処理能力が高い[16][32]。

本章の目的は、前記の心電図の非線形予測が抱えている問題に対する2VNNの改善能力を検討することにある。具体的には、IDNNと2VNNによる心電図の非線形予測学習を評価し、2VNNの学習能力を評価する。第7章では、その結果を改善された予測符号化へ用いるニューロンネットワーク予測器のニューロンネットワークを設計するための基礎データとして用いる。

## 4.2 ボルテラニューロンネットワーク

### 4.2.1 ボルテラニューロン

一対の入出力を持つ非線形ダイナミカルシステムを仮定する。便宜上、入出力は、標本化されているものとし、それぞれ $x^{(\tau)}$ ,  $z^{(\tau)}$ ,  $\tau \in \mathbf{Z}$ と表す。ここで、 $\mathbf{Z}$ は整数からなる集合を意味する。こうした入出力の関係を式(4.1), (4.2)のように表現する。式(4.1)の右辺第1項を $N$ 次離散時間ボルテラ汎関数級数と呼び、入力 $x^{(\tau)}$ と出力 $z^{(\tau)}$ の関係が連続ならば、 $N$ ,  $Q$ を十分大きくすると、ボルテラ汎関数級数[33]で表現しきれない誤差 $e^{(\tau)}$ を一様に減少でき、それらが無限大の場合、ゼロに収束することが知られている。

$$z^{(\tau)} = \sum_{n=0}^N H_n \left[ h_n; x^{(\tau)} \right] + \varepsilon^{(\tau)}, \tau \in \mathbf{Z} \quad (4.1)$$

$$H_n \left[ h_n; x^{(\tau)} \right] = \sum_{\tau_1=0}^Q \cdots \sum_{\tau_n=\tau_{n-1}}^Q h_n^{(\tau_1, \dots, \tau_n)} x^{(\tau-\tau_1)} \cdots x^{(\tau-\tau_n)} \quad (4.2)$$

ここで、 $H_n$ は $n$ 次ボルテラ汎関数、 $h_n$ は $n$ 次ボルテラ核、 $N$ はモデルの非線形次数、 $Q$ は予測次数である。

$N=1$ の場合のボルテラ汎関数モデルは、直流成分を付加した $Q$ 次線形移動平均モデルに相当する。時系列 $x(\tau)$ の予測問題は、 $y(\tau) = x(\tau+1)$ とみなしたシステム同定問題にほかならない。数学モデルとして、線形移動平均モデルを選択すれば、線形予測となる。一方、ボルテラ汎関数級数を数学モデルとして用いることにより、非線形予測が可能になる。

式(4.1)をデジタルフィルタとして具現したものをボルテラフィルタと呼ぶ。出力の線形成分は、ボルテラフィルタへの入力と1次の項の予測係数としきい値の畳み込み演算で決定される。非線形成分は、ボルテラフィルタへの入力とそれを2次、3次などの非線形項にしたものと各次の予測係数の畳み込み演算で決定される。ここで、ボルテラ汎関数級数の1次までの項と誤差を用いたものを1次ボルテラフィルタ（1<sup>st</sup>-order Volterra Filter, 以下1VFと称する）、2次までの項と誤差を用いたものを2VFと呼ぶ。2VFは、一般に内部構成が未知のシステムの入力 $x^{(t)}$ と出力 $z^{(t)}$ だけが与えられているとき、このシステムをMAモデルなどのデジタルフィルタで実現する場合に用いられるフィルタ係数を最小2乗法で求める方法を適用して用いられている。図4.1に離散型の2VFの構成、式(4.3)に2VFの入出力特性をそれぞれ示す。

$$z^{(\tau)} = \sigma_0^{(\tau)} + \sum_{p=0}^Q \sigma_1(p)^{(\tau)} x^{(\tau-p)} + \sum_{p=0}^Q \sum_{q=p}^Q \sigma_2(p, q)^{(\tau)} x^{(\tau-p)} x^{(\tau-q)} \quad (4.3)$$

ここで、 $\sigma_0$ はオフセットと離散時間2次ボルテラ汎関数級数で表現しきれない誤差、 $\sigma_1(p)$ は1番目の遅延素子入力から $Q$ 番目の遅延素子の出力までの間から得られる信号に対応する1次の項の予測係数、 $\sigma_2(p, q)$ は1番目の遅延素子の入力から $Q$ 番目の遅延素子出力までの間から得られる同一信号の組合せも含む2信号全ての組合せの積に対応する2次の項の予測係数である。

形式ニューロンの出力関数の入力部に1VFを組み込んだものを1VN、2VFを組み込んだもの

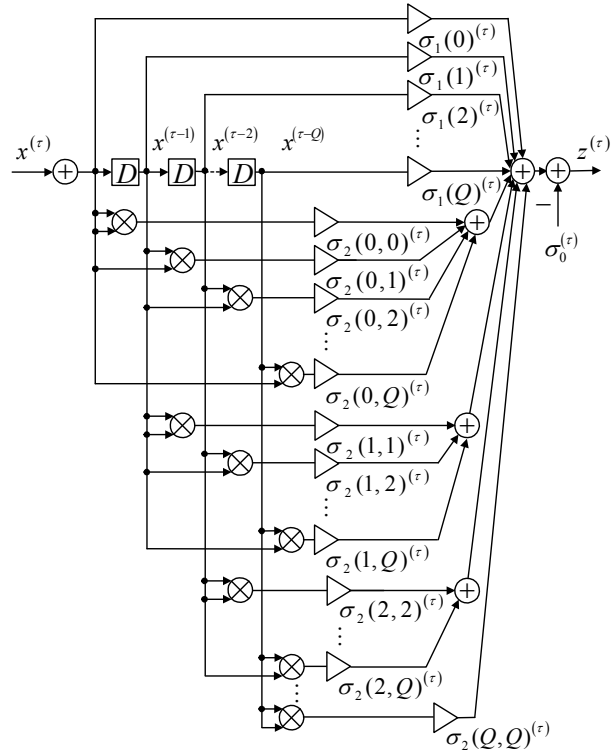


図4.1 2次ボルテラフィルタ

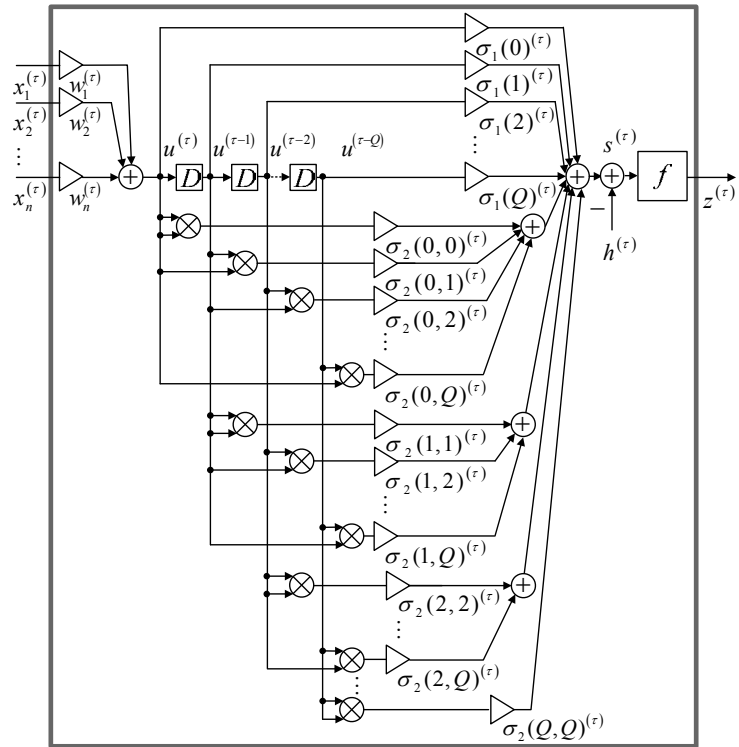


図4.2 2次ボルテラニューロン

を2VNと呼ぶ．図4.2に2VNのブロック図，式(4.4)～(4.6)に2VNの入出力特性をそれぞれ示す．

$$u^{(\tau)} = \sum_{i=1}^n w_i^{(\tau)} x_i^{(\tau)} \quad (4.4)$$

$$s^{(\tau)} = \sum_{p=0}^Q \sigma_1(p)^{(\tau)} u^{(\tau-p)} + \sum_{p=0}^Q \sum_{q=p}^Q \sigma_2(p,q)^{(\tau)} u^{(\tau-p)} u^{(\tau-q)} - h^{(\tau)} \quad (4.5)$$

$$z^{(\tau)} = f(s^{(\tau)}) = A \tan^{-1}(s^{(\tau)}) \quad (4.6)$$

ここで， $w_i$ は $i$ 番目の結合荷重， $D$ は遅延素子である．また， $w_i$ ， $h$ ， $\sigma_1(p)$ ， $\sigma_2(p, q)$ は，学習により変更される．

#### 4.2.2 入出力特性

層数が  $M$  の 2VNN の入出力特性は，入力層では式(4.7)，中間層から出力層までは式(4.8)～(4.10)となる．

$$z \left\langle \begin{smallmatrix} 1 \\ k \end{smallmatrix} \right\rangle^{(\tau)} = x \left\langle \begin{smallmatrix} 1 \\ k \end{smallmatrix} \right\rangle^{(\tau)} \quad (k=1, 2, \dots, n_1) \quad (4.7)$$

$$u \left\langle \begin{smallmatrix} L \\ k \end{smallmatrix} \right\rangle^{(\tau)} = \sum_{j=1}^{n_{L-1}} w \left\langle \begin{smallmatrix} L-1 & L \\ j & k \end{smallmatrix} \right\rangle^{(\tau)} x \left\langle \begin{smallmatrix} L-1 & L \\ j & k \end{smallmatrix} \right\rangle^{(\tau)} \\ (L=2, 3, \dots, M; j=1, 2, \dots, n_{L-1}; k=1, 2, \dots, n_L) \quad (4.8)$$

$$s \left\langle \begin{smallmatrix} L \\ k \end{smallmatrix} \right\rangle^{(\tau)} = \sum_{p=0}^Q \sigma_1 \left\langle \begin{smallmatrix} L \\ k \end{smallmatrix} \right\rangle^{(\tau)}(p) u \left\langle \begin{smallmatrix} L \\ k \end{smallmatrix} \right\rangle^{(\tau-p)} + \sum_{p=0}^Q \sum_{q=p}^Q \sigma_2 \left\langle \begin{smallmatrix} L \\ k \end{smallmatrix} \right\rangle^{(\tau)}(p, q) u \left\langle \begin{smallmatrix} L \\ k \end{smallmatrix} \right\rangle^{(\tau-p)} u \left\langle \begin{smallmatrix} L \\ k \end{smallmatrix} \right\rangle^{(\tau-q)} \\ - h \left\langle \begin{smallmatrix} L \\ k \end{smallmatrix} \right\rangle^{(\tau)} \quad (L=2, 3, \dots, M; k=1, 2, \dots, n_L) \quad (4.9)$$

$$z \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau)} = A \tan^{-1} \left( s \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau)} \right) \quad (L=2,3,\dots,M; k=1,2,\dots,n_L) \quad (4.10)$$

### 4.2.3 学習則

2VNNの評価関数は、教師信号値と第M層の出力信号値の誤差2乗和を用いて記述される式(4.11)とする。学習では、評価関数値が小さくなるように、各層において式(4.8)～(4.10)で用いられている学習可能なパラメータを変更して行く。基本となる学習則は、評価関数をこれらの各パラメータで偏微分して変更量を求める勾配法である。ここでは、学習の高速化のために、勾配法に慣性項を付加した誤差逆伝播則[34][35]を適用し、その結果、式(4.12)～(4.21)で示される2VNNの学習則が得られる。

$$E^{(\tau)} = \frac{1}{2} \sum_{k=1}^{n_M} \left( y_k^{(\tau)} - z \left\langle \begin{matrix} M \\ k \end{matrix} \right\rangle^{(\tau)} \right)^2 \quad (4.11)$$

$$w \left\langle \begin{matrix} L-1 & L \\ j & k \end{matrix} \right\rangle^{(\tau+1)} = w \left\langle \begin{matrix} L-1 & L \\ j & k \end{matrix} \right\rangle^{(\tau)} + \Delta w \left\langle \begin{matrix} L-1 & L \\ j & k \end{matrix} \right\rangle^{(\tau)} \\ (L=2,3,\dots,M; j=1,2,\dots,n_{L-1}; k=1,2,\dots,n_L) \quad (4.12)$$

$$\begin{aligned} \Delta w \left\langle \begin{matrix} L-1 & L \\ j & k \end{matrix} \right\rangle^{(\tau)} &= -\alpha \frac{\partial E^{(\tau)}}{\partial w \left\langle \begin{matrix} L-1 & L \\ j & k \end{matrix} \right\rangle^{(\tau)}} + \beta \Delta w \left\langle \begin{matrix} L-1 & L \\ j & k \end{matrix} \right\rangle^{(\tau-1)} \\ &= -\alpha r \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau)} \\ &\quad \times \left( \sigma_1(0) \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau)} + 2\sigma_2(0,0) \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau)} u \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau)} + \sum_{q=1}^Q \sigma_2(0,q) \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau)} u \left\langle \begin{matrix} L \\ k \end{matrix} \right\rangle^{(\tau-q)} \right) \\ &\quad \times z \left\langle \begin{matrix} L-1 \\ j \end{matrix} \right\rangle^{(\tau)} + \beta \Delta w \left\langle \begin{matrix} L-1 & L \\ j & k \end{matrix} \right\rangle^{(\tau-1)} \\ (L=2,3,\dots,M; j=1,2,\dots,n_{L-1}; k=1,2,\dots,n_L) \end{aligned} \quad (4.13)$$

$$h\left\langle\frac{L}{k}\right\rangle^{(\tau+1)} = h\left\langle\frac{L}{k}\right\rangle^{(\tau)} + \Delta h\left\langle\frac{L}{k}\right\rangle^{(\tau)}$$

$$(L = 2, 3, \dots, M; k = 1, 2, \dots, n_L)$$
(4.14)

$$\Delta h\left\langle\frac{L}{k}\right\rangle^{(\tau)} = -\alpha \frac{\partial E^{(\tau)}}{\partial h\left\langle\frac{L}{k}\right\rangle^{(\tau)}} + \beta \Delta h\left\langle\frac{L}{k}\right\rangle^{(\tau-1)} = \alpha r\left\langle\frac{L}{k}\right\rangle^{(\tau)} + \beta \Delta h\left\langle\frac{L}{k}\right\rangle^{(\tau-1)}$$

$$(L = 2, 3, \dots, M; k = 1, 2, \dots, n_L)$$
(4.15)

$$\sigma_1(p)\left\langle\frac{L}{k}\right\rangle^{(\tau+1)} = \sigma_1(p)\left\langle\frac{L}{k}\right\rangle^{(\tau)} + \Delta\sigma_1(p)\left\langle\frac{L}{k}\right\rangle^{(\tau)}$$

$$(L = 2, 3, \dots, M; k = 1, 2, \dots, n_L; p = 0, 1, \dots, Q)$$
(4.16)

$$\Delta\sigma_1(p)\left\langle\frac{L}{k}\right\rangle^{(\tau)} = -\alpha \frac{\partial E^{(\tau)}}{\partial \sigma_1(p)\left\langle\frac{L}{k}\right\rangle^{(\tau)}} + \beta \Delta\sigma_1(p)\left\langle\frac{L}{k}\right\rangle^{(\tau-1)}$$

$$= -\alpha r\left\langle\frac{L}{k}\right\rangle^{(\tau)} u\left\langle\frac{L}{k}\right\rangle^{(\tau-p)} + \beta \Delta\sigma_1(p)\left\langle\frac{L}{k}\right\rangle^{(\tau-1)}$$

$$(L = 2, 3, \dots, M; k = 1, 2, \dots, n_L; p = 0, 1, \dots, Q)$$
(4.17)

$$\sigma_2(p, q)\left\langle\frac{L}{k}\right\rangle^{(\tau+1)} = \sigma_2(p, q)\left\langle\frac{L}{k}\right\rangle^{(\tau)} + \Delta\sigma_2(p, q)\left\langle\frac{L}{k}\right\rangle^{(\tau)}$$

$$(L = 2, 3, \dots, M; k = 1, 2, \dots, n_L; p = 0, 1, \dots, Q; q = 0, 1, \dots, Q; p \leq q)$$
(4.18)

$$\Delta\sigma_2(p, q)\left\langle\frac{L}{k}\right\rangle^{(\tau)} = -\alpha \frac{\partial E^{(\tau)}}{\partial \sigma_2(p, q)\left\langle\frac{L}{k}\right\rangle^{(\tau)}} + \beta \Delta\sigma_2(p, q)\left\langle\frac{L}{k}\right\rangle^{(\tau-1)}$$

$$= -\alpha r\left\langle\frac{L}{k}\right\rangle^{(\tau)} u\left\langle\frac{L}{k}\right\rangle^{(\tau-p)} u\left\langle\frac{L}{k}\right\rangle^{(\tau-q)} + \beta \Delta\sigma_2(p, q)\left\langle\frac{L}{k}\right\rangle^{(\tau-1)}$$

$$(L = 2, 3, \dots, M; k = 1, 2, \dots, n_L; p = 0, 1, \dots, Q; q = 0, 1, \dots, Q; p \leq q)$$
(4.19)



$$\begin{aligned}
r \left\langle \begin{matrix} L \\ j \end{matrix} \right\rangle^{(\tau)} &= \frac{\partial E^{(\tau)}}{\partial s \left\langle \begin{matrix} L \\ j \end{matrix} \right\rangle^{(\tau)}} \\
&= \frac{A}{1 + \left( s \left\langle \begin{matrix} L \\ j \end{matrix} \right\rangle^{(\tau)} \right)^2} \sum_{k=1}^{n_{L+1}} r \left\langle \begin{matrix} L+1 \\ k \end{matrix} \right\rangle^{(\tau)} w \left\langle \begin{matrix} L & L+1 \\ j & k \end{matrix} \right\rangle^{(\tau)} \\
&\quad \times \left( \sigma_1(0) \left\langle \begin{matrix} L+1 \\ k \end{matrix} \right\rangle^{(\tau)} + 2\sigma_2(0,0) \left\langle \begin{matrix} L+1 \\ k \end{matrix} \right\rangle^{(\tau)} u \left\langle \begin{matrix} L+1 \\ k \end{matrix} \right\rangle^{(\tau)} \right. \\
&\quad \left. + \sum_{q=1}^Q \sigma_2(0,q) \left\langle \begin{matrix} L+1 \\ k \end{matrix} \right\rangle^{(\tau)} u \left\langle \begin{matrix} L+1 \\ k \end{matrix} \right\rangle^{(\tau-q)} \right) \\
&\quad (L=2,3,\dots,M-1; j=1,2,\dots,n_L; k=1,2,\dots,n_{L+1})
\end{aligned} \tag{4.20}$$

$$r \left\langle \begin{matrix} M \\ k \end{matrix} \right\rangle^{(\tau)} = \frac{\partial E^{(\tau)}}{\partial s \left\langle \begin{matrix} M \\ k \end{matrix} \right\rangle^{(\tau)}} = - \left( y_k^{(\tau)} - z \left\langle \begin{matrix} M \\ k \end{matrix} \right\rangle^{(\tau)} \right) \frac{A}{1 + \left( s \left\langle \begin{matrix} M \\ k \end{matrix} \right\rangle^{(\tau)} \right)^2} \quad (k=1,2,\dots,n_M) \tag{4.21}$$

ここで、 $\Delta\sigma_1(p)$ 、 $\Delta\sigma_2(p, q)$ はそれぞれ1次と2次の項の予測係数の変更量である。

## 4.3 計算機シミュレーション

### 4.3.1 方法

計算機シミュレーションでは、空間方向が1次元の時系列パターンの入力信号  $x(\tau)$  と教師信号  $y(\tau) = x(\tau+1)$  の組合せを学習させる。はじめに、入力遅れニューロン (Input-delay Neuron, 以下 IDN と称する) を入力層に用いて構成される3層の IDNN で学習を行い、予測誤差を評価する。図 4.3 に IDN のブロック図を示す。IDN の特徴は、形式ニューロンの入力部に有限インパルス応答 (Finite Impulse Response, 以下 FIR と称する) フィルタを装着している点である。IDN の入出力特性を式(4.22)～(4.24)に示す。

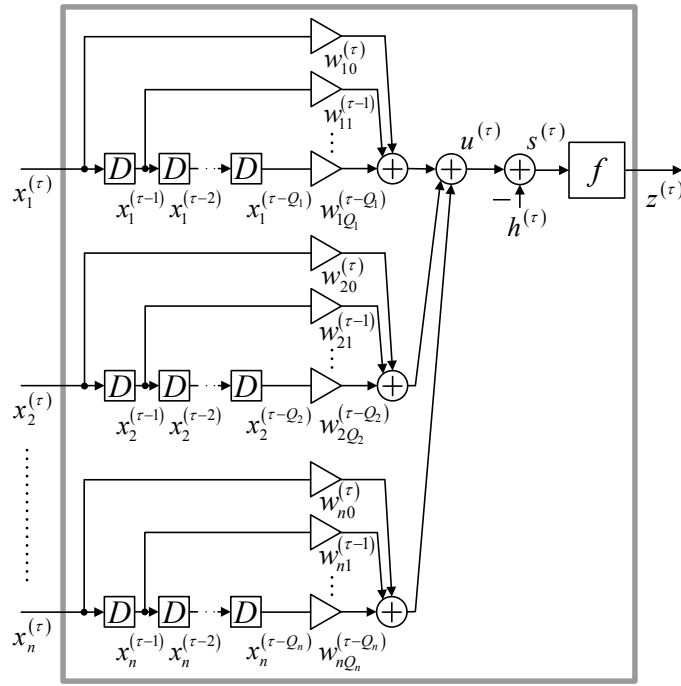


図 4.3 入力遅れニューロン

$$u^{(\tau)} = \sum_{i=1}^n \sum_{p=0}^{Q_i} w_{ip}^{(\tau)} x_i^{(\tau-p)} \quad (4.22)$$

$$s^{(\tau)} = u^{(\tau)} - h^{(\tau)} \quad (4.23)$$

$$z^{(\tau)} = f(s^{(\tau)}) = A \tan^{-1}(s^{(\tau)}) \quad (4.24)$$

ここで、 $x_i$ は $i$ 番目の入力信号、 $w_{ip}$ は $i$ 番目の入力信号に関する1番目の遅延素子の入力から $Q$ 番目の遅延素子の出力までの間から得られる信号に対応する結合荷重、 $Q_i$ は $i$ 番目の入力信号に関する遅延素子の数である。また、 $w$ 、 $h$ は、学習により変更される。図4.4にIDNNを示す。第 $i+1$ 番目の遅延素子は、第 $i$ 番目の遅延素子より1単位時間過去の信号を出力する。

次に、2VNを図4.5のように接続した1入力1出力3層の2VNNについて学習を行い、平均RMSEを比較する。学習には、健常者の安静時における正常洞調律心電図信号を用いた。その信号を図4.6に示す。この信号は、単位時間10[ms]、量子化幅0.0025[mV]の412単位時間のデータからなり、学習では、112単位時間のデータを初期入力後、入力信号と教師信号の対を1回提示する。この過程を1学習サイクルとし、これを繰り返す。表4.1に予測誤差評価計算機シミュレーション条件を示す。ここで、予測係数の初期値は、指数平滑法を用い、その他の初期値は、学習計

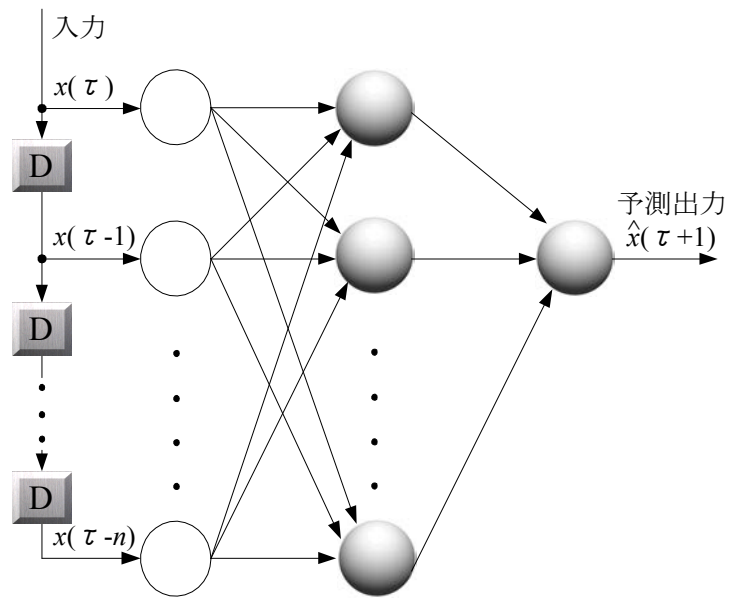


図 4.4 入力遅れニューロンによる時系列処理ネットワーク

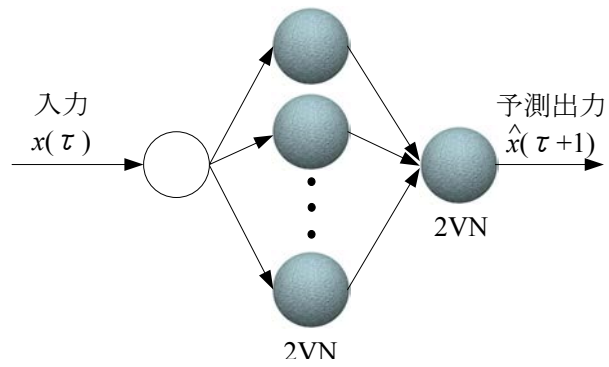


図 4.5 2 次ボルテラニューロンによる時系列処理ネットワーク

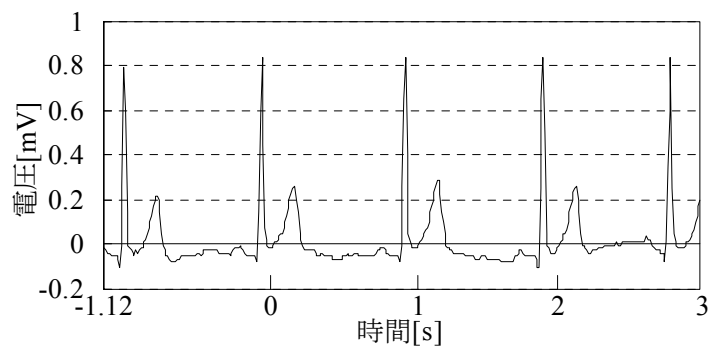


図4.6 学習用心電図信号

表4.1 予測誤差評価計算機シミュレーション条件

			IDNN	2VNN	
学習則			誤差逆伝播則	2VNNの学習則	
初期値		結合荷重	-0.3~0.3		
		しきい値	-0.3~0.3		
		予測係数	なし	$\sigma_1$	$0.7 \times 0.3^p$
$\sigma_2$	$0.7 \times 0.3^p \times 0.7 \times 0.3^q$				
中間層素子数		範囲	2~20		
		間隔	2		
タップ数, または フィルタ長		範囲	3~111	2~56	
		間隔	4	2	
学習強化係数	勾配法項	粗探索	範囲	$10^{-4} \sim 1$	
			間隔	10倍	
		細探索	範囲	$0.2 \alpha_{min} \sim 0.9 \alpha_{min}, 2 \alpha_{min} \sim 9 \alpha_{min}$	
			間隔	$0.1 \alpha_{min}, \alpha_{min}$	
	慣性項		0		
	出力係数		1		
学習サイクル		30,000			
施行回数		3			

計算機シミュレーション1回ごとに擬似乱数で決定される。また、 $\alpha_{min}$ は、粗探索学習で最小平均RMSEを得たときの勾配法項の学習強化係数である。

計算機シミュレーションは、はじめに、中間層素子数とタップ数、または、フィルタ長を計算機シミュレーション条件設定パラメータとし、それらに対して、勾配法項の学習強化係数を変化させながら、3回の粗探索学習を行い、その結果得られた平均RMSEを比較する。次に、最小平均RMSEを実現した中間層素子数を用いて、タップ数、または、フィルタ長を計算機シミュレーション条件設定パラメータとし、それらに対して、勾配法項の学習強化係数を、 $\alpha_{min}$ を元に変化させながら、3回の細探索学習を行う。その結果得られた平均RMSEを比較する。このようにして得られる両ネットワークの学習結果を用いて、それらの学習能力の評価を行う。

### 4.3.2 結果

図4.7~4.10に、計算機シミュレーションで得られた平均RMSEとその標準偏差を示す。ここで、図中の縦棒は、標準偏差の $\pm$ の範囲を示す。図4.9と図4.10より、2VNNとIDNNの各最小平均RMSEは、ほぼ同じであることが確認できる。図4.11は、IDNNと2VNNの3回の細探索学習で、最小平均RMSEが得られたときの学習開始前、学習サイクル1回目、学習サイクル100回ごとに記録した学習サイクルと1単位時間データ当りの平均評価関数値の平均との関係を示す。同図より、2VNNの1単位時間データ当りの平均評価関数値の平均は、学習サイクル15,000回

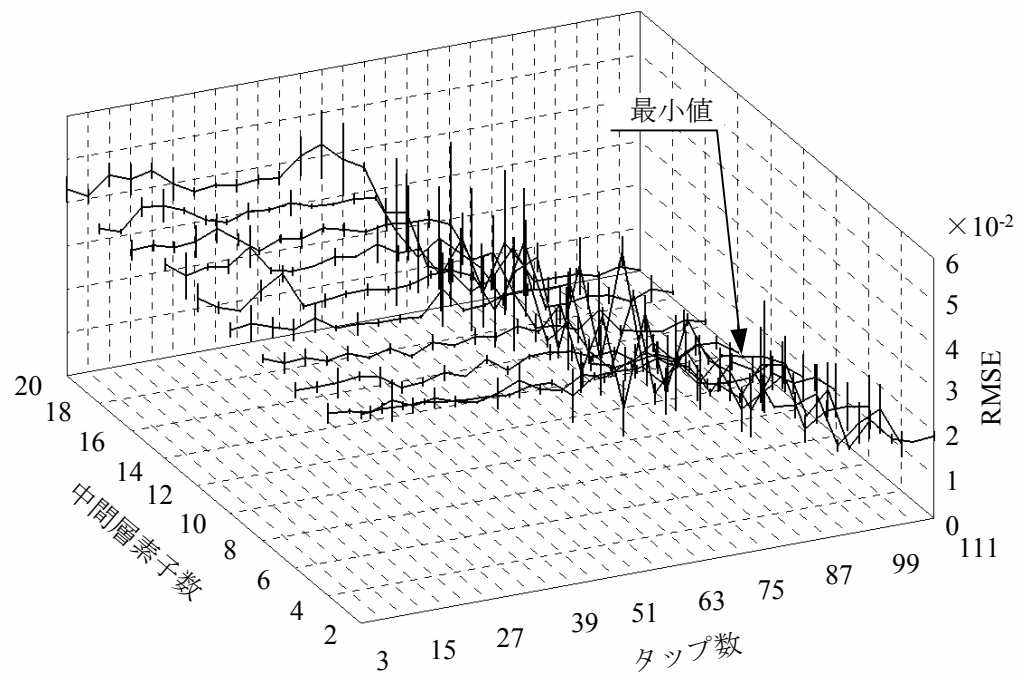


図4.7 入力遅れニューロンネットワークの粗探索学習後の平均RMSEと標準偏差

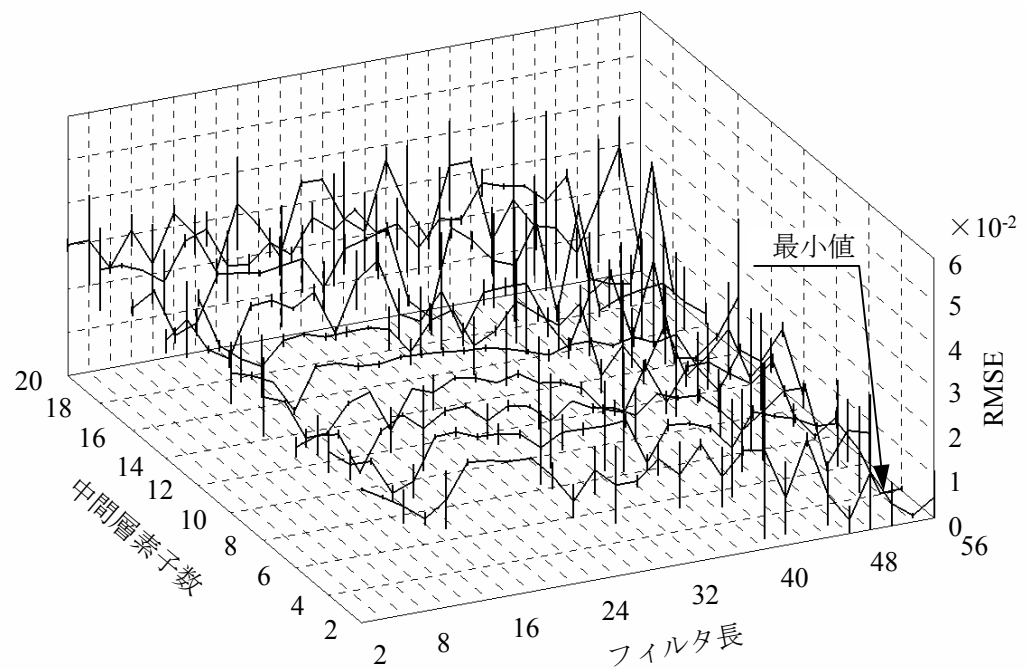


図4.8 2次ボルテラニューロンネットワークの粗探索学習後の平均RMSEと標準偏差

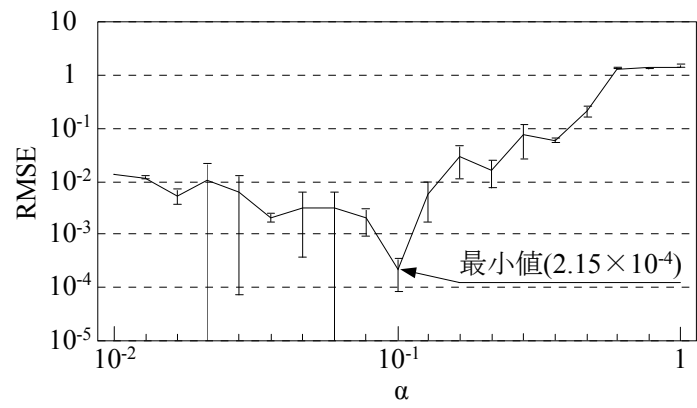


図4.9 入力遅れニューロンネットワークの細探索学習後の平均RMSEと標準偏差

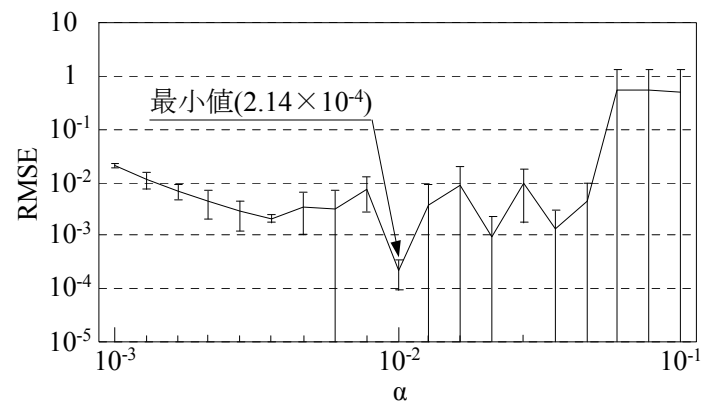


図4.10 2次ボルテラニューロンネットワークの細探索学習後の平均RMSEと標準偏差

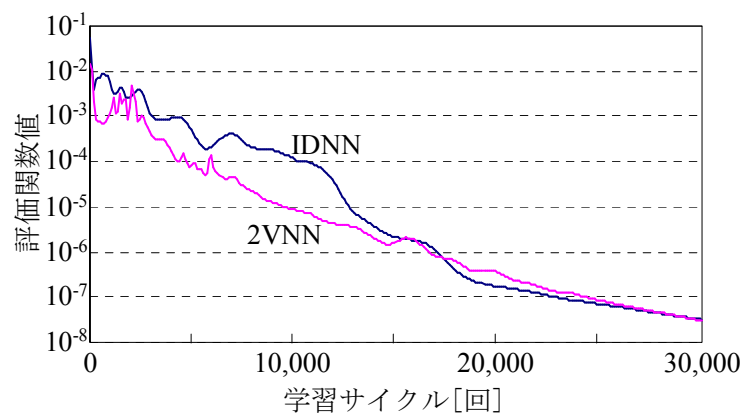


図4.11 最小平均RMSEのときの学習サイクルと1単位時間データ当りの平均評価関数値の平均との関係

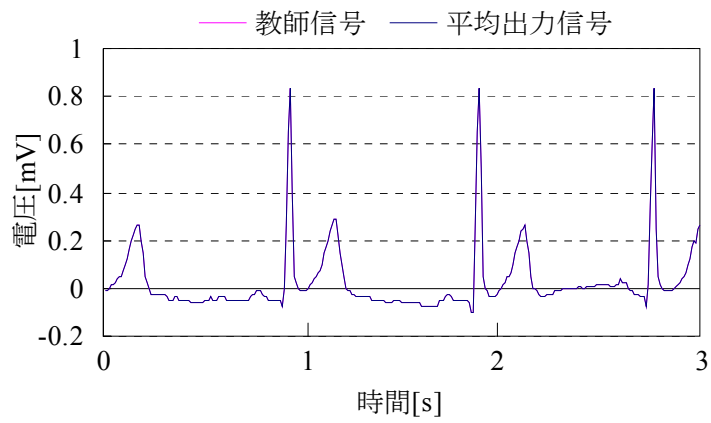


図4.12 最小平均RMSEのときの入力遅れニューロンネットワークの平均出力信号

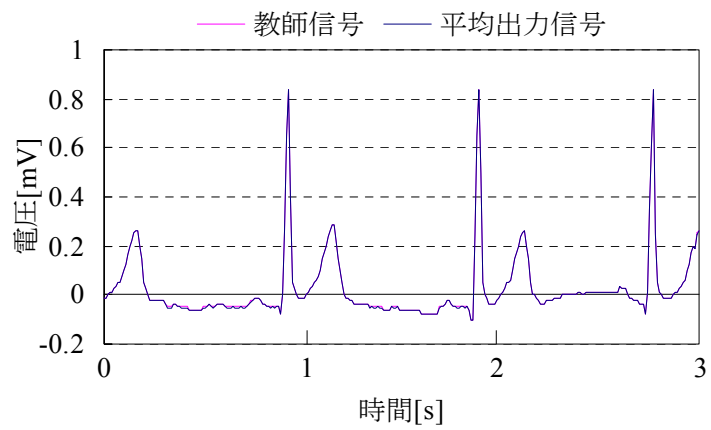


図4.13 最小平均RMSEのときの2次ボルテラニューロンネットワークの平均出力信号

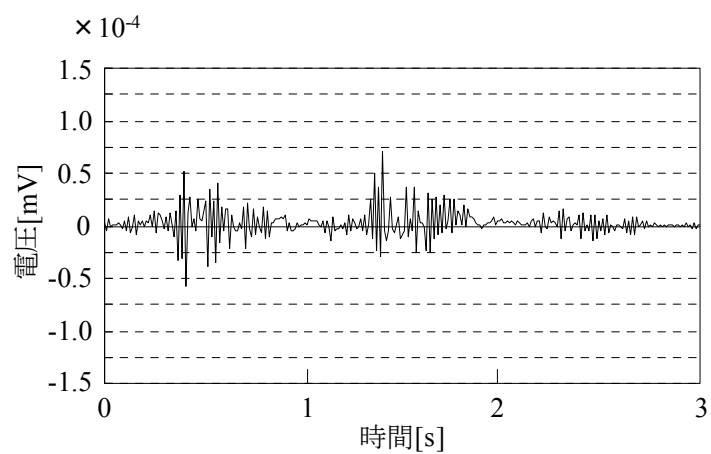


図4.14 最小平均RMSEのときの入力遅れニューロンネットワークの平均予測誤差系列

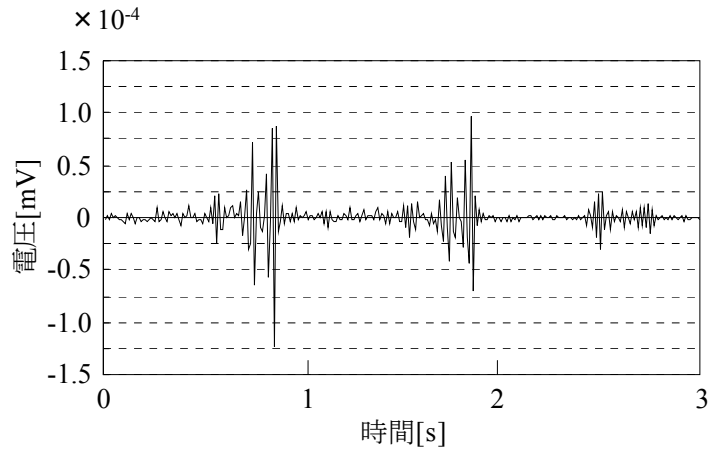


図4.15 最小平均RMSEのときの2次ボルテラニューロンネットワークの平均予測誤差系列

付近まで、IDNNより小さく、学習速度が速い。しかし、それ以降、学習サイクル30,000回付近までは、IDNNの方が僅かに小さく、学習終了時では、両ネットワークともほぼ同じになっている。図4.12と図4.13に、細探索学習で最小平均RMSEが得られたときのIDNNと2VNNの平均出力信号と教師信号を比較したものをそれぞれ示す。さらに、図4.14と図4.15に、このときのIDNNと2VNNの平均予測誤差系列をそれぞれ示す。予測誤差系列は、教師信号値から各ネットワークの平均出力信号値を引いて得た差の系列である。ここで、両ネットワークの条件は、IDNNが中間層素子数14、タップ数111、学習強化係 $1 \times 10^{-1}$ 、2VNNが中間層素子数4、フィルタ長54、学習強化係数 $1 \times 10^{-2}$ であった。図4.14と図4.15より、2VNNの平均予測誤差系列は、IDNNよりQRS波において誤差が大きい、それ以外の部分では、誤差が小さいことが示されている。

#### 4.4 考察

はじめに、予測誤差とモデルパラメータとの関係について考える。一般にネットワークのモデルパラメータ数を増やして、数学的に入出力写像の表現能力を高めても、実際には、学習能力の限界により、モデルパラメータの増加に伴う表現能力には限界が生じる。また、心電図は、各波形の生成機序が異なる時変システムであるため、ある程度タップ数やフィルタ長を増やしても、出力と因果性のない記憶データにより、入出力写像実現の妨げになる。しかし、今回のように1つの正常洞調律波形だけを学習するような場合、タップ数やフィルタ長が、その1心拍分の長さに近づくにつれて、その影響が低減されている。これは、ネットワークに1心拍



分の信号を取り込むことで、各波形の異なる生成機序を周期的な1つの平均的な生成機序、すなわち、平均的な相関関係で入出力を処理できるからではないかと考えられる。

次に、予測誤差系列について考える。心電図波形の特性として、1心拍の波形の部分ごとに生成機序が異なり、周波数成分が等しくないという点がある。心房内の興奮を反映するP波や心臓の収縮を反映するQRS波の部分は、周波数が高く、これに対して、T波、U波の部分は、周波数が低い。従来の方で心電図の予測を行った場合、QRS波での予測誤差が、ほかの部分に比較して、特に大きくなるという問題がある。これは、QRS波の生成機序に強い非線形性を含んでいるため、線形モデルでは予測しきれないことにある。2VNNは、IDNNと同等の予測誤差の改善を示しており、心電図の非線形予測に問題がないことが認められる。しかし、依然として、QRS波の予測誤差は、他の部分の波より大きいという問題が残る。この問題を改善するためにIDNNへ2VNを適用した入力遅れ2VNN（Input-delay Second-order Volterra Neuron Network, 以下ID2VNNと称する）、時間遅れニューロンネットワークへ2VNを適用した時間遅れ2VNN（Time-delay Second-order Volterra Neuron Network, 以下TD2VNNと称する）、TVNNを用いる方法が考えられる。さらに、これらをECNN[36][37]へ導入し工夫することで、モデルパラメータのみを情報源とする高圧縮率のメモリが実現できる可能性がある。

今回、2VNNとIDNNの予測誤差は、ほぼ同じであったが、実時間でのデータの記録や圧縮に対応できるようにするためには、ニューロンネットワークの計算を実時間[38]で行わなければならない。このため、高速な演算装置が必要になるという問題がある。最も計算量の多い処理は、波形の学習を繰り返す部分であり、その場合、標準的なパーソナルコンピュータのCPUの処理速度では対応できない。したがって、高速な数値演算専用のプロセッサを使用する必要がある。

プロセッサの高速化、マルチコア化、メモリの大容量化といった技術の進展は、近年著しいものがあり、従来アナログ信号として取り扱っていたデータをデジタル信号として高速処理する機器が主流となっている。今日では、マルチコア化したCPU、DSP、GPUが、様々な分野でデジタル信号処理の高速化のために応用されている。MEの分野では、大量のデータを取り扱い、高速な演算が要求される。演算装置として並列処理が可能なこれらを使用すれば、実時間処理が可能であると考えられる[9][39]。

## 4.5 結言

本論文では, IDNと形式ニューロンから構成されるIDNNと2VNから構成される2VNNの2種類の階層型ニューロンネットワークによる心電図の非線形予測学習を評価した. その結果, 両ネットワークとも, 学習精度を示す最小平均RMSEは, ほぼ同じであった.

今後の課題は, 依然として残っている比較的大きなQRS波の予測誤差を改善するために, ID2VNN, TD2VNN, TVVNNを用いて, その予測誤差の改善を確認することと, それらへECNNを適用し, その有効性を評価することである.

## 参考文献

- [1] B. Beliczynski and B. Ribeiro, "Several Enhancements to Hermite-Based Approximation of One-Variable Functions," ICANN 2008, Part I, LNCS 5163, pp.11-20, 2008.
- [2] 石川茂, 横田康成, "直交ウェーブレット変換と学習ベクトル量子化法による心電図符号化法," 信学論(D-II), J9-D-II, 9, pp.1646-1649, 1996.
- [3] S. Kadambe and P. Srinivasan, "Wavelet based neural network architecture for ECG signal compression," IX European Signal Processing Conference, pp.2437-2440, 1998.
- [4] M. V. M. Figueredo, J. C. Nievola, S. R. R. Rogal Jr. and A. B. Neto, "Compression of Electrocardiogram Using Neural Networks and Wavelets," Computer and Information Science, SCI 131, pp.27-40, 2008.
- [5] 横田康成, 岩田彰, 吉田征夫, "階層型非直交サブバンドフィルタによる心電図符号化法," 計測自動制御学会論文集, vol.32, no.11, pp.1553-1559, 1996.
- [6] G. Nave and A. Cohen, "ECG Compression Using Long-Term Prediction," IEEE Transactions on Biomedical Engineering, vol.40, no.9, pp.877-885, 1993.
- [7] Z. Arnavut, "ECG Signal Compression Based on Burrows-Wheeler Transformation and Inversion Ranks of Linear Prediction," IEEE Transactions on Biomedical Engineering, vol.54, no.3, pp.410-418, 2007.
- [8] 太田隆博, 森田啓義, "多値アルファベットに対する反辞書を用いた心電図の予測符号化," 信学技報, MBE2009-44, pp.35-40, 2009.
- [9] 長坂保典, 岩田彰, 鈴木宣夫, "ニューラルネットワークによる長時間心電図の情報圧縮," 信学論(D-II), J76-D-II, 3, pp.782-792, 1993.
- [10] Y. Nagasaka and A. Iwata, "Data Compression of Long Time ECG Recording Using BP and PCA Neural Networks," IEICE Transactions on Information and Systems, E76-D(12), pp.1434-1442, 1993.
- [11] P. Gomez-Gil and W. J. B. Oldham, "Recurrent Neural Networks as a Tool for Modeling and Prediction of Electrocardiograms," Proceedings of World Multi-Conference on Systemics, Cybernetics and Informatics, vol.4, pp.237-242, 1998.
- [12] 松岡智毅, 横田康成, "非線形予測符号化における低ビットレート域での符号化性能に関する一検討," 信学論(D-II), J82-D-II, 1, pp.100-108, 1999.

- [13] Y. Yokota, M. Koizumi, N. Matsuoka, "Prediction coding of ECG by nonlinear autoregressive model," *Systems and Computers in Japan*, vol.31, issue 7, pp.66-74, 2000.
- [14] M. A. Haque, M. K. Hasan and H. Tazawa, "Investigation of the Nonlinearity in the Heart Rate Dynamics," *Medical Engineering & Physics*, 23, pp.111-115, 2001.
- [15] M. Brito, J. Henriques, P. Gil and M. Antunes, "A Predictive Adaptive Approach to Generic ECG Data Compression," 2005 IEEE International Workshop on Intelligent Signal Processing, pp.32-37, 2005.
- [16] 小早川俊祐, 藤井崇文, 横井博一, "心電図に対するニューロンネットワークの非線形予測能力の評価," *BMFSA第20回年次大会講演論文集*, pp.9-12, 2007.
- [17] 小早川俊祐, 藤井崇文, 横井博一, "心電図に対するボルテラニューロンネットワークの非線形予測能力の評価," *第9回日本知能情報ファジィ学会九州支部学術講演会予稿集*, pp.19-22, 2007.
- [18] R. Kannan and C. Eswaran, "Lossless Compression Schemes for ECG Signals Using Neural Network Predictors," *EURASIP Journal on Advances in Signal Processing*, 35641, pp.1-20, 2007.
- [19] S. Kobayakawa, T. Fujii and H. Yokoi, "Evaluation of Prediction Capabilities of Neuron Networks Used for Electrocardiogram," *The 5th International Symposium on Management Engineering*, pp.156-161, 2008.
- [20] J. S. Taur and S. Y. Kung, "Prediction-Based Networks with ECG Application," *IEEE International Conference on Neural Networks*, pp.1920-1925, 1993.
- [21] D. S. Clouse, C. L. Giles, B. G. Horne and G. W. Cottrell, "Learning Large DeBruijn Automata with Feed-Forward Neural Networks," *Computer Science and Engineering Technical Report CS94-398*, pp.1-20, 1994.
- [22] R. El Dajani, M. Miquel, P. Rubel, P. Maison-Blanche, J. Fayn and M. Forlini, "A Neural Network Approach for Predicting and Modelling the Dynamical Behaviour of Cardiac Ventricular Repolarisation," *Proceedings of the 10<sup>th</sup> World Congress on Medical Informatics*, vol.84, pp.444-448, 2001.
- [23] R. Kannan, C. Eswaran and N. Sriraam, "Neural Network Based Methods for ECG Data Compression," *Proceedings of the 9<sup>th</sup> International Conference on Neural Information Processing*, vol.5, pp.2317-2319, 2002.
- [24] A. Chatterjee, A. Nait-Ali and P. Siarry, "An Input-delay Neural-Network-Based Approach for

- Piecewise ECG Signal Compression,” IEEE Transactions on Biomedical Engineering, vol.52, no.5, pp.945-947, 2005.
- [25] 岩本敏, 吉田達穂, 横井博一, “二足歩行ロボットのニューラル制御に関する基礎的検討,” 信学技報, MBE 93-106, pp.23-30, 1994.
- [26] 藤末義和, 猪平栄一, 横井博一, “ボルテラネットワークを用いたロボットの制御,” 信学技報, NC2003-78, pp.39-43, 2003.
- [27] 魚田紫織, 横井博一, “階層型運動スキーマによるロボットハンドの運動多様性の実現,” 信学技報, NC2003-75, pp.25-28, 2003.
- [28] S. Suematsu and H. Yokoi, “A Motion Generating System for Multi-Fingered Myoelectric Hand,” International Congress Series 1291, pp.257-260, 2006.
- [29] 三好純, 横井博一, “四輪操舵車の横滑り角学習用ニューラルネットワークの改善,” 信学技報, NC2004-107, pp.87-90, 2004.
- [30] 小早川俊祐, 横井博一, “航空機ピッチ姿勢制御のためのボルテラフィルタ内蔵型ニューラルネットワーク,” 平成17年度電気関係学会九州支部連合大会講演論文集, p.429, 2005.
- [31] 重村宗作, 西村敏博, 横井博一, “Volterra filter 組み込みNeural Network による瞬目アーチファクト除去の一手法,” 信学技報, MBE2004-87, pp.57-60, 2005.
- [32] S. Kobayakawa and H. Yokoi, “Evaluation of Prediction Capability of Non-recursion Type 2nd-order Volterra Neuron Network for Electrocardiogram,” ICONIP 2008 Neural Information Processing, pp.499-500, 2008.
- [33] 市川哲, 尺田幸男, “ボルテラ級数による高周波成分を含む発振回路の定常解析,” 信学論 (A), J-68-A, 2, pp.138-145, 1985.
- [34] D. E. Rumelhart, G. E. Hinton and R. J. Williams, “Learning representations by back-propagating errors,” Nature, vol.323, no.6088, pp.533-536, 1986.
- [35] 久間和生, 中山高, “ニューロコンピュータ工学,” 工業調査会(東京), pp.183-204, 1992.
- [36] S. Kobayakawa and H. Yokoi, “Proposal of Error Convergence-type Neuron Network System,” Extended Abstracts of 2008 International Symposium on Intelligent Informatics, p.19, 2008.
- [37] S. Kobayakawa and H. Yokoi, “Proposal of Error Convergence-type Neuron Network System,” Presented Proceeding to 2008 International Symposium on Intelligent Informatics International, Kumamoto, Japan, pp.1-10, 2008.
- [38] 内山匡, 下原勝憲, “リカレントニューラルネットワークのための実時間学習アルゴリズム

ム,” 信学論(D-II), J73-D-II, 7, pp.1158-1163, 1990.

- [39] A. Iwata, Y. Nagasaka and N. Suzumura, “Data Compression of the ECG Using Neural Network for Digital Holter Monitor,” IEEE Engineering in Medicine and Biology, vol.9, no.3, pp.53-57, 1990.

## 第5章 パラメータの時変化によるボルテラニューロンネットワークの学習精度向上

### 5.1 緒言

生体信号などの時系列信号をニューロンネットワークに学習させる場合、入力信号と教師信号との間に相関関係があっても、入力信号と教師信号との関数関係は、時々刻々微妙に変化するのが普通である。したがって、このような信号をニューロンネットワークに繰り返し提示して学習させると、入力信号と教師信号との平均的な関数関係だけが獲得されることになる。なぜなら、ニューロンネットワークの各パラメータの値は学習後に固定されるため、入力信号と教師信号との関数関係は学習後に時不変的に一意に決定されてしまうからである。ここで、ニューロンネットワークのパラメータとは、各素子の結合荷重と閾値を指す。そのため、入力信号と教師信号との時々刻々変化する関数関係を学習することができるニューロンネットワークが期待されている。

ところで、従来のニューロンネットワークには、追加学習が困難であるという問題点もある。すなわち、複数のデータを学習した後、新しいデータを学習させようとする、すでに学習したデータが損なわれ、場合によっては消去されてしまう。これを避けるためには、追加学習の度に、すでに学習したデータと新しく学習させるデータとを一緒にして学習させなければならない。このような学習方法は、学習したデータをいつまでも保存しておかなければならないだけでなく、追加学習に膨大な時間を要することになる。さらに、追加学習を何度も繰り返すと、ニューロンネットワークの学習容量をすぐに超えてしまい、それ以上学習ができなくなる。

そこで、パラメータ調整用ニューロンネットワークが提案された[1]。すなわち、最初に代表的な複数のデータを一台のニューロンネットワークに学習させておく。これを出力用ニューロンネットワークと呼ぶ。その後、最初に学習させたデータとは少し性質の異なる新しいデータを学習させる必要が生じた場合、パラメータ調整用ニューロンネットワークに学習させ、その出力信号によって、新しく学習させるデータに合うように出力用ニューロンネットワークのパラメータを微妙に調整する。この方法を用いれば、すでに学習したデータは追加学習をしても、全く損なわれず、影響を受けないことになる。また、すでに学習した結果を利用して、その上に追加学習を行うので、非常に効率的に速く追加学習ができる。さらには、出力用ニューロンネットワークのパラメータを固定して、入力信号を一つの関数にしたがって変換するので

はなく、入力信号の種類に応じてパラメータを変化させ、入力信号をさまざまな関数にしたがって変換することができるので、出力用ニューロンネットワークの学習容量は非常に大きくなる。

以上述べた追加学習用パラメータ調整用ニューロンネットワークは、入力信号と教師信号との時々刻々変化する関数関係を学習するニューロンネットワークの実現に、利用することができる。すなわち、パラメータ調整用ニューロンネットワークに学習させ、その出力信号によって、入力信号と教師信号との関数関係が時々刻々変化するよう出力用ニューロンネットワークのパラメータを微妙に調整する。これを従来の時不変ニューロンネットワークに対して、時変ニューロンネットワーク (Time Variant Neuron Network, 以下 TVNN と称する) と名づける。

本章では、出力用ニューロンネットワークに VNN を用い、そのパラメータをすべて時変化した TVVNN を提案する。TVVNN は、学習能力だけでなく、汎化能力の向上も期待される。そこで、TVVNN の入出力特性を基に学習則を導出し、TVVNN の学習精度評価法について検討する。

## 5.2 時変ボルテラニューロンネットワーク

### 5.2.1 入出力特性

第3章で述べたように、複数出力を持つニューロンネットワークに関しては、PNNを用いて対応する方が有利である。しかし、ここでは、簡易性の立場から、出力用ニューロンネットワークとパラメータ調整用ニューロンネットワークを、中間層を共通とした複数出力を持つ3層ニューロンネットワークとする。これを図5.1に示す。TVNNは、入力信号を共通とした出力用ニューロンネットワークとそのパラメータを調整するためのパラメータ調整用ニューロンネットワークから構成されている。パラメータ調整用ニューロンネットワークは、入力信号に応じて出力用ニューロンネットワークのパラメータを調整するための信号を出力する。この出力信号は、出力用ニューロンネットワークの各パラメータの値に加算される。その後、出力用ニューロンネットワークは、パラメータ調整用ニューロンネットワークに入力された入力信号と同じ入力信号に対する信号を出力する。本章では、出力用ニューロンネットワークとパラメータ調整用ニューロンネットワークに、第2層から第3層までの各素子に1VNを用いて構成されている1VNNを適用する。一般の出力用ニューロンネットワークとパラメータ調整用ニューロン



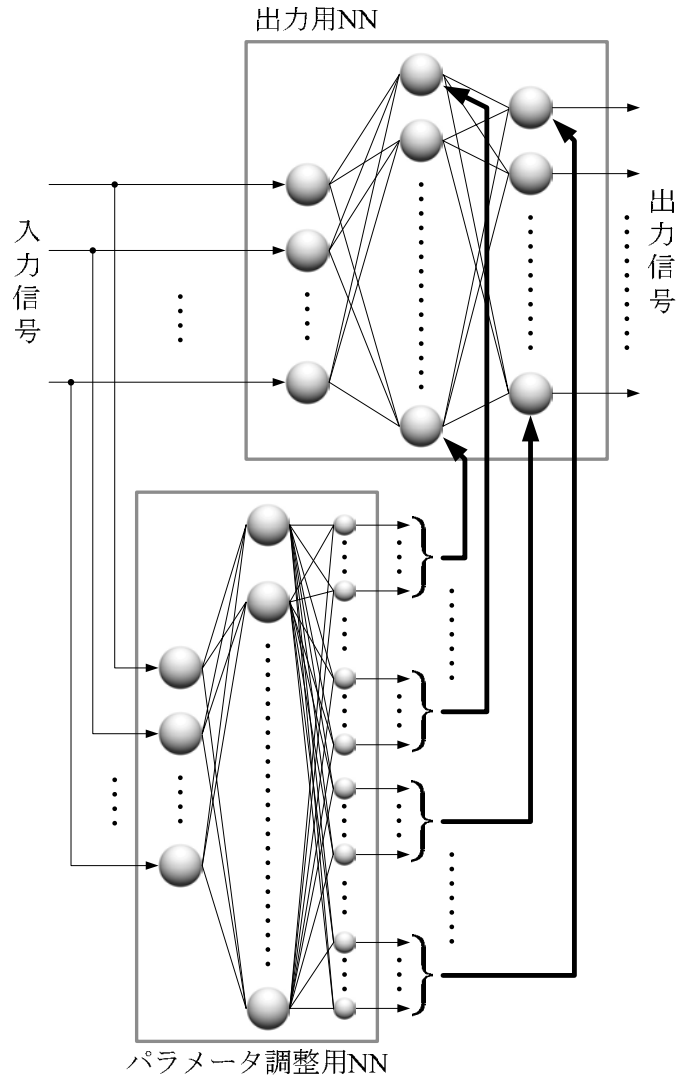


図 5.1 時変ニューロンネットワーク

ネットワークの入出力特性をそれぞれ式(5.1)～(5.4)と式(5.5)～(5.8)に示す.

$$z_o \left\langle \begin{matrix} 1 \\ i \end{matrix} \right\rangle^{(\tau)} = x_o \left\langle \begin{matrix} 1 \\ i \end{matrix} \right\rangle^{(\tau)} \quad (i=1,2,\dots,n_1) \quad (5.1)$$

$$u_o \left\langle \begin{matrix} L_o \\ k \end{matrix} \right\rangle^{(\tau)} = \sum_{j=1}^{n_{L_o-1}} \left( w_o \left\langle \begin{matrix} L_o-1 \\ j \end{matrix} \right\rangle^{(\tau)} + \Delta W \left\langle \begin{matrix} L_o-1 \\ j \end{matrix} \right\rangle^{(\tau)} \right) z_o \left\langle \begin{matrix} L_o-1 \\ j \end{matrix} \right\rangle^{(\tau)} \quad (L_o=2,3,\dots,M_o; j=1,2,\dots,n_{L_o-1}; k=1,2,\dots,n_{L_o}) \quad (5.2)$$

$$s_o \left\langle \begin{matrix} L_o \\ k \end{matrix} \right\rangle^{(\tau)} = \sum_{p=0}^Q \left( \sigma_{o1}(p) \left\langle \begin{matrix} L_o \\ k \end{matrix} \right\rangle^{(\tau)} + \Delta \Sigma_1(p) \left\langle \begin{matrix} L_o \\ k \end{matrix} \right\rangle^{(\tau)} \right) u_o \left\langle \begin{matrix} L_o \\ k \end{matrix} \right\rangle^{(\tau-p)} - \left( h_o \left\langle \begin{matrix} L_o \\ k \end{matrix} \right\rangle^{(\tau)} + \Delta H \left\langle \begin{matrix} L_o \\ k \end{matrix} \right\rangle^{(\tau)} \right)$$

$$(L_o = 2, 3, \dots, M_o; k = 1, 2, \dots, n_{L_o}) \quad (5.3)$$

$$z_o \left\langle \frac{L_o}{k} \right\rangle^{(\tau)} = A_o \tan^{-1} \left( \left\langle \frac{L_o}{k} \right\rangle^{(\tau)} \right) \\ (L_o = 2, 3, \dots, M_o; k = 1, 2, \dots, n_{L_o}) \quad (5.4)$$

$$z_v \left\langle \frac{1}{i} \right\rangle^{(\tau)} = x_v \left\langle \frac{1}{i} \right\rangle^{(\tau)} \quad (i = 1, 2, \dots, n_1) \quad (5.5)$$

$$u_v \left\langle \frac{L_v}{m} \right\rangle^{(\tau)} = \sum_{l=1}^{n_{L_v-1}} w_v \left\langle \frac{L_v-1}{l} \frac{L_v}{m} \right\rangle^{(\tau)} z_v \left\langle \frac{L_v-1}{l} \right\rangle^{(\tau)} \\ (L_v = 2, 3, \dots, M_v; l = 1, 2, \dots, n_{L_v-1}; m = 1, 2, \dots, n_{L_v}) \quad (5.6)$$

$$s_v \left\langle \frac{L_v}{m} \right\rangle^{(\tau)} = \sum_{p=0}^Q \sigma_{v1}(p) \left\langle \frac{L_v}{m} \right\rangle^{(\tau)} u_v \left\langle \frac{L_v}{m} \right\rangle^{(\tau-p)} - h_v \left\langle \frac{L_v}{m} \right\rangle^{(\tau)} \\ (L_v = 2, 3, \dots, M_v; m = 1, 2, \dots, n_{L_v}) \quad (5.7)$$

$$z_v \left\langle \frac{L_v}{m} \right\rangle^{(\tau)} = A_v \tan^{-1} \left( \left\langle \frac{L_v}{m} \right\rangle^{(\tau)} \right) \\ (L_v = 2, 3, \dots, M_v; m = 1, 2, \dots, n_{L_v}) \quad (5.8)$$

ここで、各記号の添え字の $o$ は出力用ニューロンネットワーク、 $v$ はパラメータ調整用ニューロンネットワークのものを示す。 $\Delta W$ ,  $\Delta H$ ,  $\Delta \Sigma_1$ は、それぞれパラメータ調整用ニューロンネットワークから出力される出力用ニューロンネットワークの結合荷重、しきい値、1次ボルテラ係数の各調整量を示す。

## 5.2.2 学習則

TVNNの学習は、出力用ニューロンネットワークとパラメータ調整用ニューロンネットワークを別々に行う。はじめに、出力用ニューロンネットワークをパラメータ調整用ニューロンネットワークから独立させた状態でTVNNに与えられた教師信号に対して学習を行う。次に、学習済みの出力用ニューロンネットワークとパラメータ調整用ニューロンネットワークを接続

し、出力用ニューロンネットワークのパラメータの値を固定させた状態で、パラメータ調整用ニューロンネットワークを先の出力用ニューロンネットワークの学習時に用いたTVNNに与えられた教師信号に対して学習を行う．すなわち、学習時の誤差は、出力用ニューロンネットワークを経て、パラメータ調整用ニューロンネットワークに伝播される．実際の学習では、評価関数値が小さくなるように、出力用ニューロンネットワークとパラメータ調整用ニューロンネットワークの各層において、それぞれ式(5.2), (5.3)と式(5.6), (5.7)で用いられている学習可能なパラメータを変更して行く．ここでは、誤差逆伝播則を適用し、その結果、式(5.9)～(5.21)で示される1VNNを用いたTVVNNの学習則が得られる．ここで、TVVNNの評価関数は、TVVNNに与えられた教師信号値とその第3層の出力信号値の誤差2乗和を用いて記述される式(5.9)とする．

$$E = \frac{1}{2} \sum_{k=1}^{n_3} \left( y \left\langle \frac{3}{k} \right\rangle^{(\tau)} - z_o \left\langle \frac{3}{k} \right\rangle^{(\tau)} \right)^2 \quad (5.9)$$

$$\begin{aligned} \Delta w_o \left\langle \frac{L_o-1}{j} \frac{L_o}{k} \right\rangle^{(\tau)} &= -\alpha_o \frac{\partial E^{(\tau)}}{\partial w_o \left\langle \frac{L_o-1}{j} \frac{L_o}{k} \right\rangle^{(\tau)}} + \beta_o \Delta w_o \left\langle \frac{L_o-1}{j} \frac{L_o}{k} \right\rangle^{(\tau-1)} \\ &= -\alpha_o r_o \left\langle \frac{L_o}{k} \right\rangle^{(\tau)} \sigma_{o1}(0) \left\langle \frac{L_o}{k} \right\rangle^{(\tau)} z_o \left\langle \frac{L_o-1}{j} \right\rangle^{(\tau)} + \beta_o \Delta w_o \left\langle \frac{L_o-1}{j} \frac{L_o}{k} \right\rangle^{(\tau-1)} \\ &\quad (L_o = 2, 3, \dots, M_o; j = 1, 2, \dots, n_{L_o-1}; k = 1, 2, \dots, n_{L_o}) \end{aligned} \quad (5.10)$$

$$\begin{aligned} \Delta h_o \left\langle \frac{L_o}{k} \right\rangle^{(\tau)} &= -\alpha_o \frac{\partial E^{(\tau)}}{\partial h_o \left\langle \frac{L_o}{k} \right\rangle^{(\tau)}} + \beta_o \Delta h_o \left\langle \frac{L_o}{k} \right\rangle^{(\tau-1)} = \alpha_o r_o \left\langle \frac{L_o}{k} \right\rangle^{(\tau)} + \beta_o \Delta h_o \left\langle \frac{L_o}{k} \right\rangle^{(\tau-1)} \\ &\quad (L_o = 2, 3, \dots, M_o; k = 1, 2, \dots, n_{L_o}) \end{aligned} \quad (5.11)$$

$$\begin{aligned} \Delta \sigma_1(p)_o \left\langle \frac{L_o}{k} \right\rangle^{(\tau)} &= -\alpha_o \frac{\partial E^{(\tau)}}{\partial \sigma_1(p)_o \left\langle \frac{L_o}{k} \right\rangle^{(\tau)}} + \beta_o \Delta \sigma_{o1}(p) \left\langle \frac{L_o}{k} \right\rangle^{(\tau-1)} \\ &= -\alpha_o r_o \left\langle \frac{L_o}{k} \right\rangle^{(\tau)} u_o \left\langle \frac{L_o}{k} \right\rangle^{(\tau-p)} + \beta_o \Delta \sigma_{o1}(p) \left\langle \frac{L_o}{k} \right\rangle^{(\tau-1)} \\ &\quad (L_o = 2, 3, \dots, M_o; k = 1, 2, \dots, n_{L_o}; p = 0, 1, \dots, Q) \end{aligned} \quad (5.12)$$

$$\begin{aligned}
r_o \left\langle \begin{matrix} L_o \\ j \end{matrix} \right\rangle^{(\tau)} &= \frac{\partial E^{(\tau)}}{\partial s_o \left\langle \begin{matrix} L_o \\ j \end{matrix} \right\rangle^{(\tau)}} \\
&= \frac{A_o}{1 + \left( s_o \left\langle \begin{matrix} L_o \\ j \end{matrix} \right\rangle^{(\tau)} \right)^2} \sum_{k=1}^{n_{L_o+1}} r_o \left\langle \begin{matrix} L_o+1 \\ k \end{matrix} \right\rangle^{(\tau)} w_o \left\langle \begin{matrix} L_o & L_o+1 \\ j & k \end{matrix} \right\rangle^{(\tau)} \sigma_{ol}(0) \left\langle \begin{matrix} L_o+1 \\ k \end{matrix} \right\rangle^{(\tau)} \\
&\quad (L_o = 2, 3, \dots, M_o - 1; j = 1, 2, \dots, n_{L_o}; k = 1, 2, \dots, n_{L_o+1})
\end{aligned} \tag{5.13}$$

$$\begin{aligned}
r_o \left\langle \begin{matrix} M_o \\ k \end{matrix} \right\rangle^{(\tau)} &= \frac{\partial E^{(\tau)}}{\partial s_o \left\langle \begin{matrix} M_o \\ k \end{matrix} \right\rangle^{(\tau)}} = - \left( y_k^{(\tau)} - z_o \left\langle \begin{matrix} M_o \\ k \end{matrix} \right\rangle^{(\tau)} \right) \frac{A_o}{1 + \left( s_o \left\langle \begin{matrix} M_o \\ k \end{matrix} \right\rangle^{(\tau)} \right)^2} \\
&\quad (k = 1, 2, \dots, n_M)
\end{aligned} \tag{5.14}$$

$$\begin{aligned}
\Delta w_v \left\langle \begin{matrix} L_v-1 & L_v \\ l & m \end{matrix} \right\rangle^{(\tau)} &= -\alpha_v \frac{\partial E^{(\tau)}}{\partial w_v \left\langle \begin{matrix} L_v-1 & L_v \\ l & m \end{matrix} \right\rangle^{(\tau)}} + \beta_v \Delta w_v \left\langle \begin{matrix} L_v-1 & L_v \\ l & m \end{matrix} \right\rangle^{(\tau-1)} \\
&= -\alpha_v r_v \left\langle \begin{matrix} L_v \\ m \end{matrix} \right\rangle^{(\tau)} \sigma_{vl}(0) \left\langle \begin{matrix} L_v \\ m \end{matrix} \right\rangle^{(\tau)} z_v \left\langle \begin{matrix} L_v-1 \\ l \end{matrix} \right\rangle^{(\tau)} + \beta_v \Delta w_v \left\langle \begin{matrix} L_v-1 & L_v \\ m & l \end{matrix} \right\rangle^{(\tau-1)} \\
&\quad (L_v = 2, 3, \dots, M_v; l = 1, 2, \dots, n_{L_v-1}; m = 1, 2, \dots, n_{L_v})
\end{aligned} \tag{5.15}$$

$$\begin{aligned}
\Delta h_v \left\langle \begin{matrix} L_v \\ m \end{matrix} \right\rangle^{(\tau)} &= -\alpha_v \frac{\partial E^{(\tau)}}{\partial h_v \left\langle \begin{matrix} L_v \\ m \end{matrix} \right\rangle^{(\tau)}} + \beta_v \Delta h_v \left\langle \begin{matrix} L_v \\ m \end{matrix} \right\rangle^{(\tau-1)} = \alpha_v r_v \left\langle \begin{matrix} L_v \\ m \end{matrix} \right\rangle^{(\tau)} + \beta_v \Delta h_v \left\langle \begin{matrix} L_v \\ m \end{matrix} \right\rangle^{(\tau-1)} \\
&\quad (L_v = 2, 3, \dots, M_v; m = 1, 2, \dots, n_{L_v})
\end{aligned} \tag{5.16}$$

$$\begin{aligned}
\Delta \sigma_1(p)_v \left\langle \begin{matrix} L_v \\ m \end{matrix} \right\rangle^{(\tau)} &= -\alpha_v \frac{\partial E^{(\tau)}}{\partial \sigma_{v1}(p) \left\langle \begin{matrix} L_v \\ m \end{matrix} \right\rangle^{(\tau)}} + \beta_v \Delta \sigma_{v1}(p) \left\langle \begin{matrix} L_v \\ m \end{matrix} \right\rangle^{(\tau-1)} \\
&= -\alpha_v r_v \left\langle \begin{matrix} L_v \\ m \end{matrix} \right\rangle^{(\tau)} u_v \left\langle \begin{matrix} L_v \\ m \end{matrix} \right\rangle^{(\tau-p)} + \beta_v \Delta \sigma_{v1}(p) \left\langle \begin{matrix} L_v \\ m \end{matrix} \right\rangle^{(\tau-1)} \\
&\quad (L_v = 2, 3, \dots, M_v; m = 1, 2, \dots, n_{L_v}; p = 0, 1, \dots, Q)
\end{aligned} \tag{5.17}$$

$$r_v \left\langle \begin{matrix} L_v \\ l \end{matrix} \right\rangle^{(\tau)} = \frac{\partial E^{(\tau)}}{\partial s_v \left\langle \begin{matrix} L_v \\ l \end{matrix} \right\rangle^{(\tau)}}$$

$$\begin{aligned}
&= \frac{A_v}{1 + \left( s_v \left\langle \frac{L_v}{l} \right\rangle^{(\tau)} \right)^2} \sum_{m=1}^{n_{L_v+1}} r_v \left\langle \frac{L_v+1}{m} \right\rangle^{(\tau)} w_v \left\langle \frac{L_v}{l} \quad \frac{L_v+1}{m} \right\rangle^{(\tau)} \sigma_{vl}(0) \left\langle \frac{L_v+1}{m} \right\rangle^{(\tau)} \\
&\left( L_v = 2, 3, \dots, M_v - 1; l = 1, 2, \dots, n_{L_v}; m = 1, 2, \dots, n_{L_v+1} \right)
\end{aligned} \tag{5.18}$$

$$\begin{aligned}
r_v \left\langle \frac{M_v}{m} \right\rangle^{(\tau)} &= \frac{\partial E^{(\tau)}}{\partial s_v \left\langle \frac{M_v}{m} \right\rangle^{(\tau)}} \\
&= \frac{A_v}{1 + \left( s_v \left\langle \frac{M_v}{m} \right\rangle^{(\tau)} \right)^2} r_o \left\langle \frac{L_o}{k} \right\rangle^{(\tau)} \left( \sigma_{ol}(0) \left\langle \frac{L_o}{k} \right\rangle^{(\tau)} + \Delta \Sigma_l(0) \left\langle \frac{L_o}{k} \right\rangle^{(\tau)} \right) z_o \left\langle \frac{L_o-1}{j} \right\rangle^{(\tau)} \\
&\left( L_o = 2, 3, \dots, M_o; j = 1, 2, \dots, n_{L_o-1}; k = 1, 2, \dots, n_{L_o}; m = 1, 2, \dots, n_{M_v} \right)
\end{aligned} \tag{5.19}$$

$$\begin{aligned}
r_v \left\langle \frac{M_v}{m} \right\rangle^{(\tau)} &= \frac{\partial E^{(\tau)}}{\partial s_v \left\langle \frac{M_v}{m} \right\rangle^{(\tau)}} = - \frac{A_v}{1 + \left( s_v \left\langle \frac{M_v}{m} \right\rangle^{(\tau)} \right)^2} r_o \left\langle \frac{L_o}{k} \right\rangle^{(\tau)} \\
&\left( L_o = 2, 3, \dots, M_o; m = 1, 2, \dots, n_{M_v} \right)
\end{aligned} \tag{5.20}$$

$$\begin{aligned}
r_v \left\langle \frac{M_v}{m} \right\rangle^{(\tau)} &= \frac{\partial E^{(\tau)}}{\partial s_v \left\langle \frac{M_v}{m} \right\rangle^{(\tau)}} = \frac{A_v}{1 + \left( s_v \left\langle \frac{M_v}{m} \right\rangle^{(\tau)} \right)^2} r_o \left\langle \frac{L_o}{k} \right\rangle^{(\tau)} u_o \left\langle \frac{L_o}{k} \right\rangle^{(\tau)} \\
&\left( L_o = 2, 3, \dots, M_o; k = 1, 2, \dots, n_{L_o}; m = 1, 2, \dots, n_{M_v} \right)
\end{aligned} \tag{5.21}$$

ここで、式(5.19)～(5.21)は、それぞれ出力用ニューロンネットワークの結合荷重調整量、しきい値調整量、1次ボルテラ係数調整量を出力するパラメータ調整用の出力層のニューロンの強化信号である。

### 5.3 学習精度の評価法

TVVNNの汎化能力を評価するために、過去の心電図非線形予測評価で得られた1VNNの緒元を基に、その遅延素子数とフィルタ長を標本周波数の倍率に合わせた別の1VNNで構成されるTVVNNと心電図信号の生成機序の時変性が強い不整脈心電図を時変現象が十分含まれる長さ

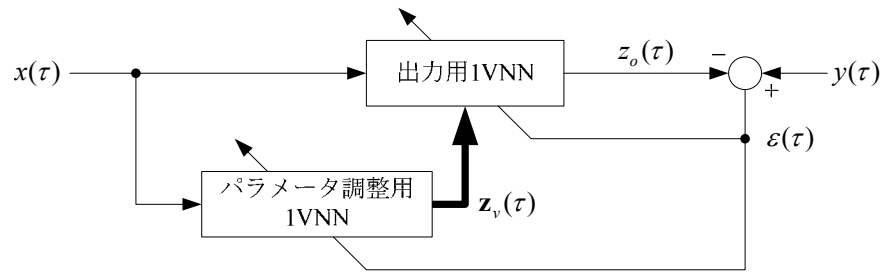


図5.3 心電図予測用時変ボルテラニューロンネットワーク

で用いて、心電図予測に関する学習計算機シミュレーションを行う。心電図予測用TVVNNを図5.3に示す。図中の $\mathbf{z}_v$ は、パラメータ調整用1VNNの出力信号ベクトルである。

はじめに、学習用心電図信号の最大絶対値が1となるようにゲインチューニングする。単独にした出力用1VNNへ空間方向が1次元の時系列パターンを入力信号 $x(\tau)$ と教師信号 $y(\tau)=x(\tau+1)$ の組合せを与え学習強化係数を変えながら探索学習を行う。ここで、最小平均RMSEを得た条件の中から最小RMSEの1VNNを選び、それを出力用1VNNとする。次に、パラメータ調整用1VNNへ入力信号 $x(\tau)$ を入力し、決定した出力用ニューロンネットワークへ教師信号 $y(\tau)=x(\tau+1)$ を与え学習強化係数を変えながら探索学習を行う。ここで、TVVNNが最小平均RMSEを得た条件の

の中から最小RMSEの1VNNを選び、それをパラメータ調整用ニューロンネットワークとする。

学習は、出力用ニューロンネットワークとパラメータ調整用ニューロンネットワークの各最大記憶データ長の長い方のデータ長分の信号を初期入力後、入力信号と教師信号の対を1回提示する。この過程を1学習サイクルとし、これを繰り返す。計算機シミュレーションは、勾配法項の学習強化係数を変化させながら、3回の探索学習を行い、その結果得られた平均RMSEを比較する。ここで、予測係数の初期値は、指数平滑法を用いる。その他の初期値は、学習計算機シミュレーション1回ごとに、擬似乱数で決定される。

以上の探索学習で得られたTVVNNへ学習後の不整脈心電図信号を初期入力し、その後、連続する未学習不整脈心電図信号を入力して、その汎化精度を評価する。

## 5.4 考察

TVNNの学習能力について考察する。出力用ニューロンネットワークは、一般のニューロン

ネットワーク単体と同じである。このニューロンネットワークの学習能力には限界がある。しかし、入力信号と教師信号の関数関係が時変の場合、数学的にパラメータ調整用ニューロンネットワークを用いて、さらに出力用ニューロンネットワークが学習しきれない誤差の学習を行うことができる。これにより、学習能力の改善ができる場合があると考えられる。この学習能力の改善効果を出力用ニューロンネットワーク単体と TVNN の学習後の学習精度を比較し、評価しなければならない。しかし、パラメータ調整用ニューロンネットワークは、一般に出力用ニューロンネットワークの学習誤差を完全になくすことを保証していない。この問題を解決する方法に関しては、次章で述べる。

次に、TVNN の汎化能力について考察する。出力用ニューロンネットワークは、入力信号と出力信号の変化する関数関係の平均的な関数を学習で獲得する。パラメータ調整用ニューロンネットワークは、出力用ニューロンネットワークのその学習で獲得した平均的な関数で表現できない部分の関数を出力用ニューロンネットワークで表現できるようにする関数を学習する。しかし、この関数も入力信号の変化に対して一定ではなく、さらに変化するものであれば、パラメータ調整用ニューロンネットワークも、その平均的な関数を学習で獲得するにすぎない。これでは、汎化の際、十分に能力は発揮できないと考えられる。このような場合は、さらに、高次のパラメータ調整用ニューロンネットワークが考えられる。しかし、学習を行っていない関数関係に関しては、TVNN の汎化能力を発揮できない。このため、効率よく未学習関数を学習するためのシステムやアルゴリズムが必要であると考えられる。

本章では、TVVNN の学習精度と汎化精度を、正常洞調律心電図信号を用いて評価を行い、心電図予測に関しての TVVNN の基本的な学習能力や汎化能力の検証と評価ができるものと考えられる。今後の課題は、実際に学習計算機シミュレーションを行い、それらの能力の検証と評価をすることである。

## 5.5 結言

本章では、学習計算機シミュレーション時間が比較的少なく、かつ、非線形予測能力がある IVNN を用いて、TVVNN を構成し、その有効性を評価する方法を検討した。その結果、学習能力に関しては、入力信号と教師信号の関数関係が時変の場合、数学的にパラメータ調整用ニューロンネットワークを用いて、さらに出力用ニューロンネットワークが学習しきれない誤差の学習を行うことができると考えられることから、その改善ができると考えられる。しかし、パ

ラメータ調整用ニューロンネットワークは、一般に出力用ニューロンネットワークの学習誤差を完全になくすことを保証していない。この問題を解決する方法に関しては、次章で述べる。汎化能力に関しては、パラメータ調整用ニューロンネットワークを用いて向上させることが可能であると考えられるが、その入力信号と教師信号の関数関係が一定ではないならば、さらに高次のパラメータ調整用ニューロンネットワークが必要と考えられる。しかし、学習を行っていない関数関係に関しては、汎化能力が期待できないため、効率よく未学習関数関係を学習するためのシステムやアルゴリズムが必要であると考えられる。今後の課題は、実際に正常洞調律心電図信号を用いて学習計算機シミュレーションを行い、心電図予測に関してそれらの能力の検証と評価をすることである。



## 参考文献

- [1] E. Inohira, H. Oonishi and H. Yokoi, “Perturbational Neural Networks for Incremental Learning in Virtual Learning System,” ICONIP 2007, Part I, Lecture Notes in Computer Science, Springer, vol.4984, pp.395-404, 2008.

## 第6章 誤差収束の原理

### 6.1 緒言

ニューロンネットワークは、入力信号と教師信号の間に相関関係がなければ、学習することはできない。しかし、入力信号と教師信号との間には、相関のある成分と無相関の成分が混在しているのが普通である。したがって、入力信号と教師信号の間の無相関の成分については、ニューロンネットワークで学習できないことになる。入力信号と教師信号の間の相関のある成分についても、ニューロンネットワークの学習能力が低い場合には、十分に学習できない。この相関成分を完全に学習させるためには、ニューロンネットワークの学習能力を向上させる必要がある。さらに、入力信号と教師信号との関数関係が、一意に決まらないようにパラメータの時変化も必要になる。しかし、このような高い学習能力を持ったニューロンネットワークは、まだ存在していない。

現在まで、ニューロンネットワークの学習精度を向上させるために、盛んに研究が行われてきた。構成素子であるニューロンの入出力特性やニューロンネットワークの構造を改善するなど、多くの方法が提案されており、学習精度向上に有効であることが示されている。しかし、これらの方法では、学習回数をいくら多くしても、ニューロンネットワークの出力誤差をゼロに収束させることは保証されていない。したがって、ニューロンネットワークは、許容範囲内の出力誤差を含んだ状態で使用されているのが現状である。しかし、出力誤差が小さければ小さいほど良い予測符号化[1],[2]や航空宇宙機の飛行制御[3],[4]などにニューロンネットワークを応用する場合には、現在のニューロンネットワークの学習精度では不十分である。

本章では、学習精度が低い学習システムでも、それを複数段用いることにより、出力誤差を限りなくゼロに近づけるための原理を提案する。すなわち、入力信号と教師信号の間に一定の関数関係がある場合には、その関数関係について、1段目の学習システムを学習し、その近似関数を獲得する。2段目の入力信号と1段目の出力誤差信号の関数関係について、2段目の学習システムを学習し、その近似関数を獲得する。3段目の入力信号と2段目の出力誤差信号の関数関係について、3段目の学習システムを学習し、その近似関数を獲得する。これを何段も繰り返す。こうして、出力誤差は徐々にゼロに収束していく。入力信号と教師信号の間に一定の関数関係がなく、相関関係だけがある場合には、平均的な関数関係が学習されていく。したがって、段数を増やしても、出力誤差をゼロに近づけてゆくことはできない。このような場合、パ

ラメータを時変化すれば、出力誤差をゼロに近づけてゆくことは可能である。なお、入力信号と教師信号の間に相関があるだけではなく、教師信号系列、あるいは、出力誤差信号系列において、時間の前後で相関関係があれば、学習システムにそれらの関係も学習させることで、出力誤差をさらに速くゼロに近づけてゆくことも可能である。

以上述べた誤差収束の原理に基づいて、学習システムを工学的に実現しようとする、いくつかの問題が生じる。例えば、雑音の存在や段数が増えるとともに、教師信号が微弱になってしまうなどの問題である。本章では、それらの問題点を解決できる誤差収束型学習システムの工学的実現法についても提案する。最後に、誤差収束型学習システムの応用例について、いくつかの提案を行う。

## 6.2 原理

図6.1は、誤差収束の原理をブロック図で示したものである。まず、入力信号 $x$ と教師信号 $y$ の間に一定の関数関係がある場合について考察する。各段の学習システムの入力信号 $x$ は同一である。第1段目の学習システムでは、入力信号 $x$ と教師信号 $y_1$  ( $y_1=y$ ) との関数関係を学習するが、学習回数をどんなに増やしても完全には学習できないので、その近似関数を獲得することになる。したがって、第1段目の学習システムからは教師信号 $y_1$ と出力信号 $z_1$ の差である出力誤差 $\varepsilon_1$ が生じる。次に、第2段目の学習システムでは、第1段目の出力誤差 $\varepsilon_1$ を教師信号 $y_2$ とし、入力信号 $x$ との関数関係を学習する。この場合も完全な学習はできないので、その近似関数を獲得することになる。したがって、第2段目の学習システムからも教師信号 $y_2$ と出力信号 $z_2$ の差である出力誤差 $\varepsilon_2$ が生じることになる。しかし、この出力誤差は段数を重ねるごとに徐々に小さくなっていく。

一方、学習システム全体の出力信号 $z$ は、各段の学習システムの出力信号 $z_1, z_2, \dots$ の総和となる。各段で学習が完了するたびに、出力信号が加えられて、出力信号 $z$ が教師信号 $y$ に近づいてゆく。以上の説明を式(6.1)～(6.4)に示す。

$$y_i = z_i + \varepsilon_i \quad (i=1,2,\dots,n) \quad (6.1)$$

$$z_i = f_i(x) \quad (i=1,2,\dots,n) \quad (6.2)$$

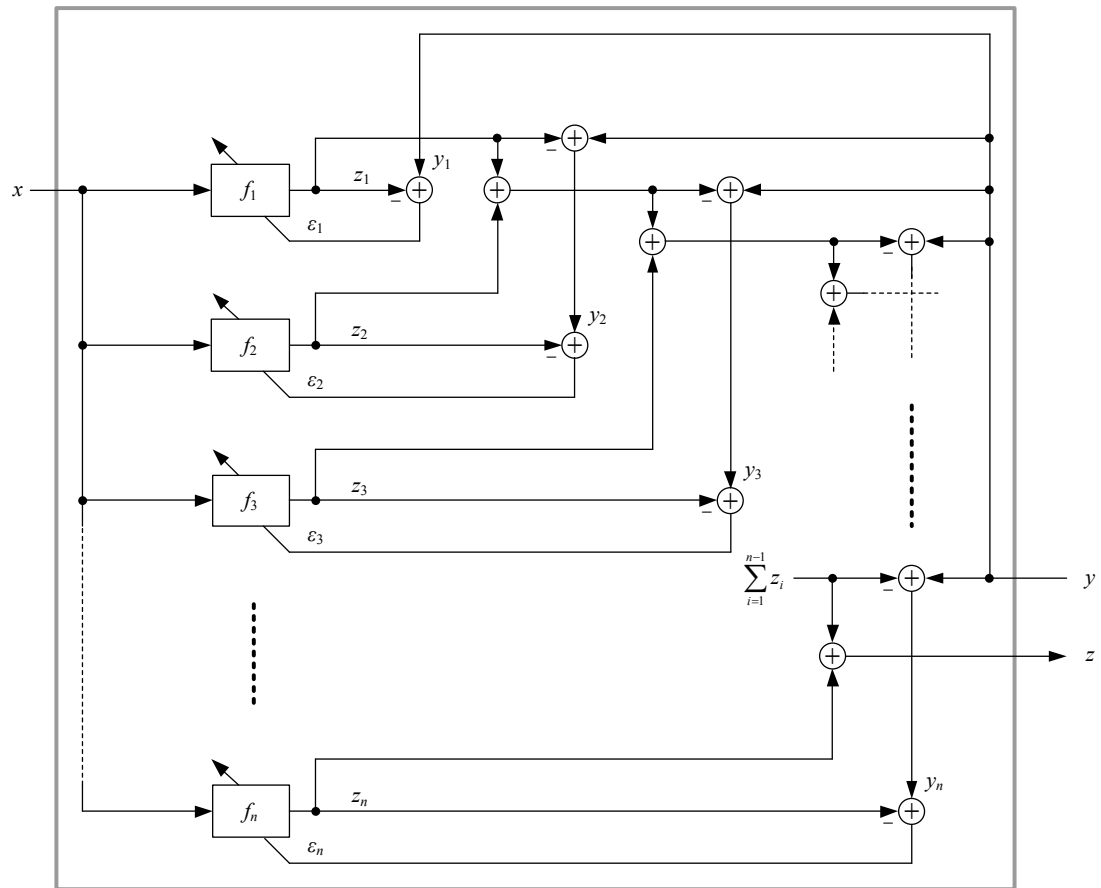


図 6.1 誤差収束の原理のブロック図

$$\varepsilon_i = y_{i+1} \quad (i=1,2,\dots,n) \quad (6.3)$$

$$y = \sum_{i=1}^n z_i + \varepsilon_n \quad (6.4)$$

ここで、 $f_i$ は第  $i$  段目の入出力関係の近似関数である。また、入出力信号数を任意とする場合は、各信号を成分とするスカラー信号やベクトル信号の組合せで表現できる。

### 6.3 工学的実現

誤差収束の原理を実際に適用する場合、式(6.4)の誤差項をゼロに収束させることができない。実際の工学処理では、信号は量子化され、制限された有効数字の値として取り扱われる。誤差項は、学習可能な近似関数で次々と表していくと次第に小さくなり、学習が困難になる問題が

生じる．このため，誤差項を増幅し，それを学習して近似関数を作成する．この操作で新たに発生する誤差は，増幅した誤差項の増幅率が 1 より大きい場合，増幅前の信号レベルに復元されると，操作前の誤差より小さくなる．さらに，その増幅率が非常に大きい場合，復元時の誤差は，ゼロに近づく．ここで， $n$  番目の近似関数で表現する誤差の増幅率を  $A_n$  として近似した結果得られた誤差が， $\varepsilon_{An}$  になると仮定する．このとき，原出力に対する出力誤差  $\varepsilon_n$  は，式(6.5)で示され，式(6.6)に示す関係があるとする．したがって，近似関数の段数が高くなるに従って， $\varepsilon_n$  は，次第に小さくなる．また， $\varepsilon_n$  は，近似関数の復元後の出力の総和を出力信号値の有効数字に合わせて端数処理を行うことで，ゼロになる．これを式(6.7)に示す．以上より，前記問題を改善するためのこの手段は，高精度出力を得るために効果的かつ有用であることが示された．これを量子化信号の誤差収束の原理と呼ぶ．

$$\varepsilon_n = \frac{\varepsilon_{An}}{A_n} \quad (6.5)$$

$$|\varepsilon_n| < |\varepsilon_{An}| \quad (6.6)$$

$$\begin{aligned} y &= f_r \left( \sum_{i=1}^{n_{ef}} \frac{z_i}{A_i} \right) + \varepsilon_{n_{ef}} \\ &= f_r \left( \sum_{i=1}^{n_{ef}} \frac{z_i}{A_i} \right) \end{aligned} \quad (6.7)$$

ここで， $f_r$  は出力信号値の有効数字に合わせた端数処理関数， $n_{ef}$  は誤差をゼロにするために必要な近似関数の段数である．

次に，量子化信号の誤差収束の原理を基に設計した ECFS を提案する．この処理概念とブロック図をそれぞれ図 6.2 と図 6.3 に示す．ECFS は，内部の関数の入出力関係に応じて，最適な処理ができる関数と学習則を自由に選択して構成することができる．ここでは，ECFS の一般的な入出力特性を具体的に ECFS へ適用した関数の入出力特性に触れずに述べる．

ECFS は，各段で関数の誤差を収束させるためと適切な学習を行うために用いる入力信号と教師信号の振幅を調整する増幅器を備えている．さらに，ECFS は，各段の関数の出力信号を ECFS の出力信号レベルへ復元し，誤差を収束させるために，それらの出力部に教師信号の増

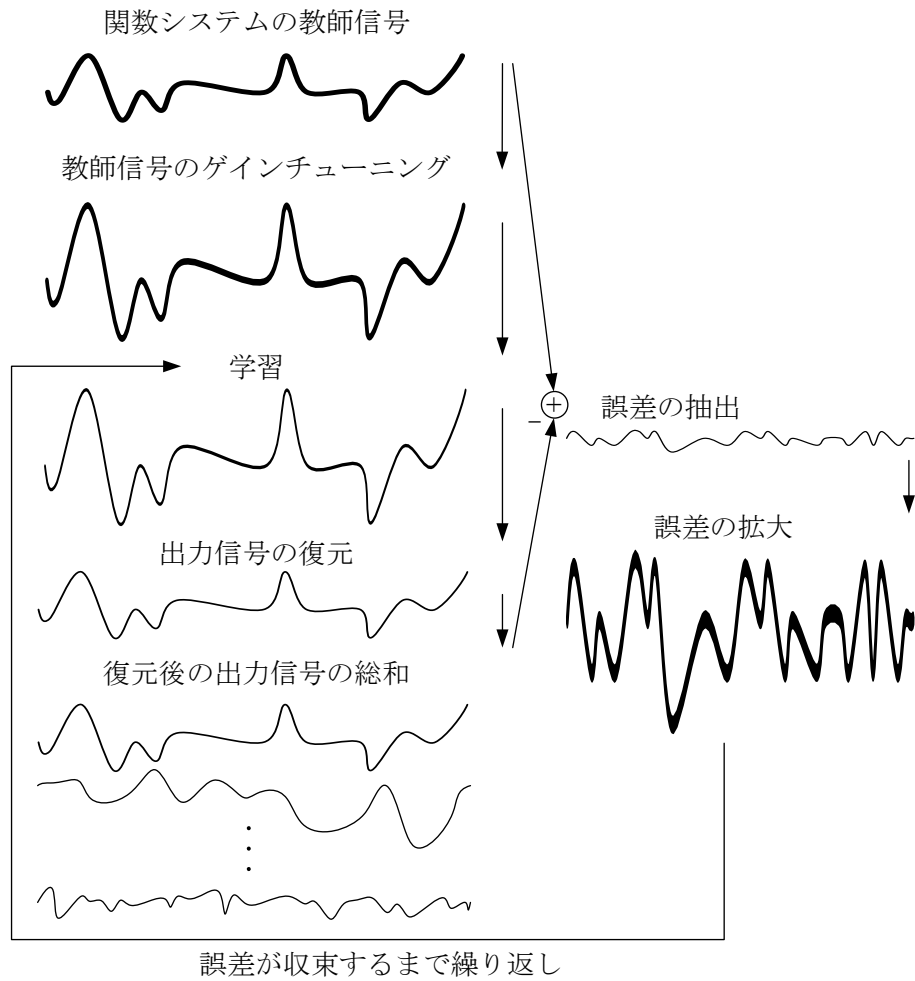


図 6.2 誤差収束関数システムの処理概念

幅率の逆数を増幅率とする増幅器を備えている．ECFS の入出力特性を式(6.8)～(6.11)に示す．さらに，教師信号の関係を式(6.12)～(6.14)に示す．ここで，式(6.11)～(6.13)で用いられる増幅率の条件を式(6.15)に示す．

$$x_i = A_{ini} x \quad (i=1, 2, \dots, n) \quad (6.8)$$

$$z_{Ai} = f_i(x_i) \quad (i=1, 2, \dots, n) \quad (6.9)$$

$$z = f_r \left( \sum_{i=1}^n z_i \right) \quad (6.10)$$

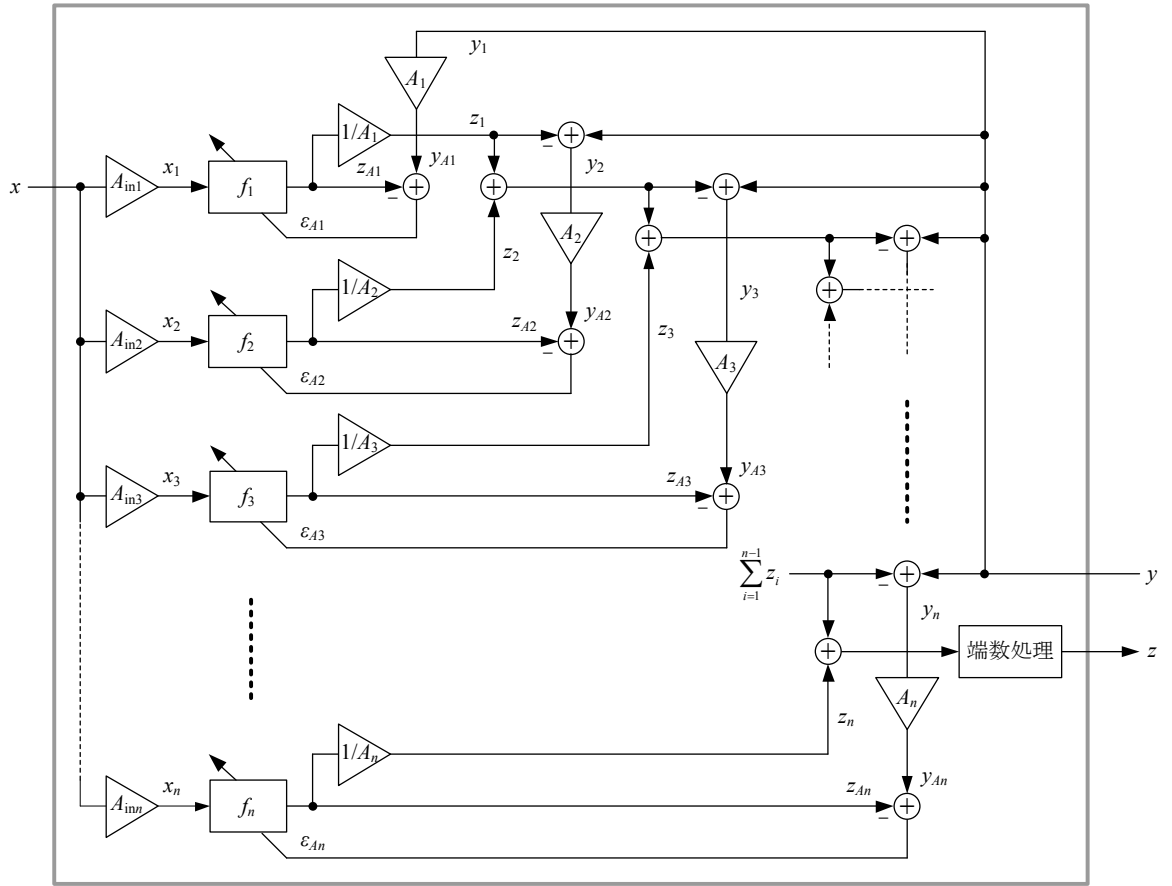


図 6.3 誤差収束関数システム

$$z_i = \frac{z_{Ai}}{A_i} \quad (i=1,2,\dots,n) \quad (6.11)$$

$$y_{A1} = A_1 y \quad (6.12)$$

$$y_{Ai} = A_i \left( y - \sum_{j=1}^{i-1} z_j \right) \quad (i=2,3,\dots,n) \quad (6.13)$$

$$\varepsilon_{Ai} = y_{Ai} - z_{Ai} \quad (i=1,2,\dots,n) \quad (6.14)$$

$$A_i \neq 0 \quad (i=1,2,\dots,n) \quad (6.15)$$

ここで、 $x_i$ は第  $i$  段目の関数の増幅後の入力信号、 $A_{ini}$ は第  $i$  段目の関数の入力信号増幅率、 $z_{Ai}$

は第  $i$  段目の関数の出力信号,  $y$  は ECFS の教師信号,  $z$  は ECFS の出力信号,  $y_i$  は第  $i$  段目の関数の増幅前の教師信号,  $y_{A1}$  と  $y_{Ai}$  はそれぞれ第 1 段目と第  $i$  段目の関数の増幅後の教師信号,  $z_i$  は第  $i$  段目の関数の復元後の出力信号,  $A_i$  は第  $i$  段目の関数の教師信号増幅率,  $\varepsilon_{Ai}$  は第  $i$  段目の関数の増幅後の教師信号に対する出力誤差を示す.

次に, 複数出力の関数システムへ ECFS を適用する方法について述べる. 複数出力の関数システムの出力数と同じ数の ECFS を用意し, それらを並列ユニットとした並列システムの形で, 複数出力の関数システムを表現する. これを ECPFS と呼ぶ. 図 6.4 に ECPFS を示す. また, ECPFS の入出力特性を式(6.16)~(6.23)に示す. さらに, 教師信号の関係を式(6.24)~(6.26)に示す. ここで, 式 (6.23)~(6.25)で用いられる増幅率の条件は, 式(6.27)となる.

$$n_{in} \geq 1 \quad (n_{in} \in \mathbb{Z}) \quad (6.16)$$

$$\mathbf{x} = (x_1, x_2, \dots, x_{n_{in}}) \quad (6.17)$$

$$\mathbf{A}_{inij} = (a_{ij1}, a_{ij2}, \dots, a_{ijn_{in}}) \quad (i=1, 2, \dots, m; j=1, 2, \dots, n_i) \quad (6.18)$$

$$x_{Aijk} = a_{ijk} x_k \quad \left( \begin{array}{l} i=1, 2, \dots, m; j=1, 2, \dots, n_i \\ k=1, 2, \dots, n_{in} \end{array} \right) \quad (6.19)$$

$$\mathbf{x}_{ij} = (x_{Aij1}, x_{Aij2}, \dots, x_{Aijn_{in}}) \quad (i=1, 2, \dots, m; j=1, 2, \dots, n_i) \quad (6.20)$$

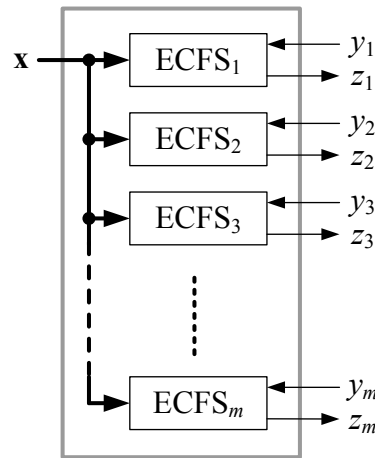


図 6.4 誤差収束並列関数システム



$$z_{Aij} = f_{ij}(\mathbf{x}_{ij}) \quad (i=1,2,\dots,m; j=1,2,\dots,n_i) \quad (6.21)$$

$$z_i = f_r\left(\sum_{j=1}^{n_i} z_{ij}\right) \quad (i=1,2,\dots,m) \quad (6.22)$$

$$z_{ij} = \frac{z_{Aij}}{A_{ij}} \quad (i=1,2,\dots,m; j=1,2,\dots,n_i) \quad (6.23)$$

$$y_{Ail} = A_{il}y_i \quad (i=1,2,\dots,m) \quad (6.24)$$

$$y_{Aij} = A_{ij}\left(y_i - \sum_{k=1}^{j-1} z_{ik}\right) \quad (i=1,2,\dots,m; j=2,3,\dots,n_i) \quad (6.25)$$

$$\varepsilon_{Aij} = y_{Aij} - z_{Aij} \quad (i=1,2,\dots,m; j=1,2,\dots,n_i) \quad (6.26)$$

$$A_{ij} \neq 0 \quad (i=1,2,\dots,m; j=1,2,\dots,n_i) \quad (6.27)$$

ここで、 $n_{in}$  は入力信号数、 $\mathbf{x}$  は入力信号の集合、 $x_A$  は増幅後の入力信号、 $\mathbf{A}_{in}$  は入力信号増幅率の集合、 $a$  は入力信号増幅率、 $x$  と  $x_{Aij}$  と  $a_{ij}$  の添え字は信号番号、2つの数字の組合せで表記されている添え字の1番目の数字は並列ユニットの番号、2番目の数字はECFSの関数の段番号、 $y_i$  は第 $i$ 番目の並列ユニットの教師信号、 $z_i$  は第 $i$ 番目の並列ユニットの出力信号を示す。

## 6.4 応用例

ここで、ECFS と ECPFS を用いた2つの応用システムを提案する。例えば、プラントのシミュレータを実現させる場合、入力信号として状態信号と制御信号が必要である。さらに、入力信号に従って決定される時々刻々のシミュレータの状態出力信号は、状態入力信号として再帰的に必要である。高精度シミュレータを実現させる場合、その出力に誤差が含まれない方がよい。したがって、指定された有効数字で動作するシミュレータは、ECFS とその出力誤差をな

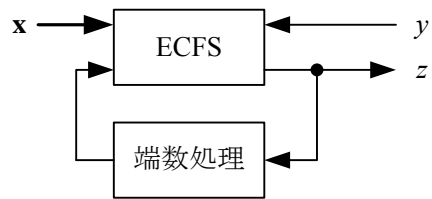


図 6.5 誤差収束リカレントシステム

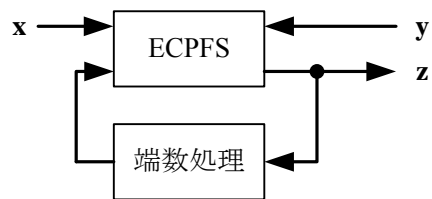


図 6.6 誤差収束並列リカレントシステム

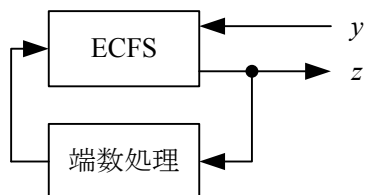


図 6.7 系列信号逐次出力用誤差収束リカレントシステム

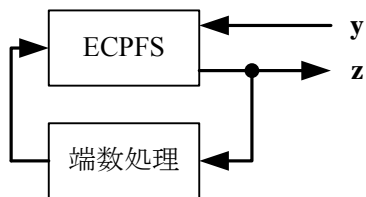


図 6.8 系列信号逐次出力用誤差収束並列リカレントシステム

くするための端数処理を用いたプラントモデルとすることで、高精度な出力が得られると考えられる．具体的には、1 状態信号のシミュレータとして、図 6.5 に示すような ECFS の出力信号を、端数処理を介して入力に帰還させたシステムが考えられる．これを ECRS と呼ぶ．さらに、複数状態信号のシミュレータとして、図 6.5 に示すような ECPFS の出力信号ベクトルを、端数処理を介して入力に帰還させたシステムが考えられる．これを ECPRS と呼ぶ．これらの図の  $\mathbf{x}$  は入力信号ベクトル、 $y$  は教師信号、 $\mathbf{y}$  は教師信号ベクトル、 $z$  は出力信号、 $\mathbf{z}$  は出力信号ベクトルである．また、ECRS と ECPRS のほかの使用例として、図 6.5 と図 6.6 の入力信号ベクトル  $\mathbf{x}$  を取り除いき、初期入力を与えるだけで、系列信号の逐次出力が得られるようにした図 6.7 と図 6.8 に示すリカレントシステムが考えられる．以上までの各システムに用いる学習可能な関数として、ニューロンネットワークを適用することができる．特に ECFS と ECPFS の関数にニューロンネットワークを適用したものをそれぞれ ECNN、ECPNN と呼ぶ．

## 6.5 考察

はじめに、ECFSの出力の高精度化と学習の高速化について考察する．ECFSの出力誤差は、複数の関数によって収束するため、少ない学習回数で高精度出力が得られる．ここで、ECFSの関数の段数を無限に設定した場合、最終段の関数が学習で使用する前段の関数の誤差信号は、非常に小さくなる．すなわち、最終段の関数の教師信号の増幅率は非常に大きくなり、ECFSの出力信号レベルへ復元したときの出力誤差が、ECFSに与えられた教師信号と比較して皆無に等しくなるため、すべての関数の出力信号の総和は、非常に高精度な出力信号になる．ECFSの学習サイクルは、この条件下で1個の教師データだけを学習するならば、理論的に1回となる[5]．例えば、学習則に確率的降下法を用いる場合、前記に示したように高精度な学習が1データ毎実行されるため、安定した出力が数回の学習サイクルで得られると期待される．また、ECFSの出力信号に対する各段の関数の出力信号の成分比率は、学習サイクルの増加とともに、より前段で大きくなると考えられる．前記の効果は、関数の段数の増加とともに顕著になり、学習精度や学習速度を向上させると考えられる．

次に、ECFSの学習能力について考察する．離散時間型システムで使用する信号の標本周波数を高くすれば、その能力は向上する．しかし、デジタル回路の動作周波数、メモリ容量、データ処理時間の制約から、標本周波数の限界が生じる．さらに、データ処理が効率的に行われるための標本周波数は、標本化定理によって、その最低周波数が決定される．しかし、そのデ

ータを用いて離散時間型システムの良好な学習が得られない場合、標本周波数を高くすることで、データ数を増やした状態での学習を行わなければならない。これにより、データ処理が不効率となる。ECFSは、標本周波数を高くせずに、この問題を解決できると考えられる。

実際には、ECNNを目的とする写像に対応できるように工夫しながら具現化して研究を行う。そして、学習能力などの評価を行ってゆく。

## 6.6 結言

本章では、相関関係がある入出力を学習可能な近似関数とその近似関数によって生じる誤差を次々と表す複数の別の学習可能な近似関数で表し、誤差を収束させる誤差収束の原理を考案した。さらに、この原理の工学的な実現方法として、相関関係がある量子化入出力信号を学習可能な近似関数とその近似関数によって生じる誤差を次々と表す別の学習可能な近似関数で表し、それらの近似関数の出力信号の総和を量子化出力信号の有効数字の桁数に合わせて端数処理を行うことで、誤差をゼロにする量子化信号の誤差収束の原理を考案した。また、この量子化信号の誤差収束の原理を具現したものとして、単出力で用いる ECFS と複出力で用いる ECPFS を考案した。さらに、これらを用いた応用システムとして、非線形プラントのシミュレータや系列信号逐次出力器として使うことが期待できる ECRS と ECPRS も考案した。

以上、ECFS は、出力の高精度化、学習の高速化、離散時間システムの学習能力改善に関して理論的に優れており、将来のプラントモデリングのための主導的な基礎システムとして、十分にその能力を持っていると述べることができる。今後の課題は、ECFS で用いる関数にニューロンネットワークを適用した ECNN を、対象とする入出力関係に対応できるように工夫しながら具現し、その学習能力の評価を行うことである。

## 参考文献

- [1] 小早川俊祐, 藤井崇文, 横井博一, “心電図に対するニューロンネットワークの非線形予測能力の評価,” BMFSA 第 20 回年次大会講演論文集, pp.9-12, 2007.
- [2] S. Kobayakawa, T. Fujii and H. Yokoi, “Evaluation of Prediction Capabilities of Neuron Networks Used for Electrocardiogram,” Proceedings of the 5th International Symposium on Management Engineering, Kitakyushu, Japan, pp.156-161, 2008.
- [3] 小早川俊祐, 横井博一, “航空機ピッチ姿勢制御のためのボルテラフィルタ内蔵型ニューラルネットワーク,” 平成 17 年度電気関係学会九州支部連合大会講演論文集, p.429, 2005.
- [4] 小早川俊祐, 横井博一, “並列化ニューロンネットワークの航空機における予測問題への適用,” 信学技報, vol.106, no.471, SANE2006-119-133, pp.43-45, 2007.
- [5] 小早川俊祐, 横井博一, “誤差収束型ニューロンネットワークシステムの収束能力に関する基礎的実験,” 第 8 回情報科学技術フォーラム講演論文集, 2 分冊, pp.531-537, 2009.

## 第7章 誤差収束によるニューロンネットワークシステムの学習精度向上

### 7.1 緒言

第2章で、一般的に時系列信号を学習する1入力1出力のBPNは、同じ入力信号値に対して、複数の教師信号値が対応することがあることを述べた。このような状態では、BPNは学習することができず、ECNNのニューロンネットワークとして用いることができない。この問題を解決するために、IDNNやVNNを用いる手段がある。これらの手段は、予測符号化を用いた非線形信号の圧縮についての従来の研究[1],[2]に用いられていた。非線形予測のためのニューロンネットワークの学習は、過去信号と予測信号の相関性が強いほど学習しやすい。したがって、ECNNの第1段目のニューロンネットワークにおける学習は、比較的やさしい。しかし、高い段数のニューロンネットワークの学習は、入出力信号の関数関係が、段数が上がるに連れて弱くなるため、難しくなる。

本章の目的は、前記の学習の困難性を改善できる ECNN を用いた予測器を提案することにある。はじめに、ECNN の学習能力を改善するために、その再設計を行う。その再設計は、ECNN の各段のニューロンネットワークが、入力信号と教師信号の関数関係が強い予測器となるように行われる。その結果、ECNNP を考案した。また、ECNNP に 2VNN, 再帰型 2 次 VNN (Recurrent Second-order Volterra Neuron Network, 以下 R2VNN と称する), それらの入力部に FIR フィルタを装着した ID2VNN や入力遅れ R2VNN (Input-delay Recurrent Second-order Volterra Neuron Network, 以下 IDR2VNN と称する) を用いることを提案する。ここでは、正常洞調律心電図信号を用いて、2 段の 2VNN から構成される ECNNP の学習計算機シミュレーションを行う。そして、ECNNP の学習精度を評価し、誤差のない学習、すなわち、完全学習を実証する。これにより、系列信号逐次出力器が実現され、高精度予測器の実現も期待できることを述べる。さらに、この ECNNP の緒元を基に構成した別の 2 段の 2VNN で構成される ECNNP と条件の異なる正常洞調律心電図信号を用いて、その学習能力を評価する。最後に、その学習能力の限界を、ID2VNN で構成される ECNNP で改善できることを示す。

### 7.2 誤差収束型ニューロンネットワーク予測器

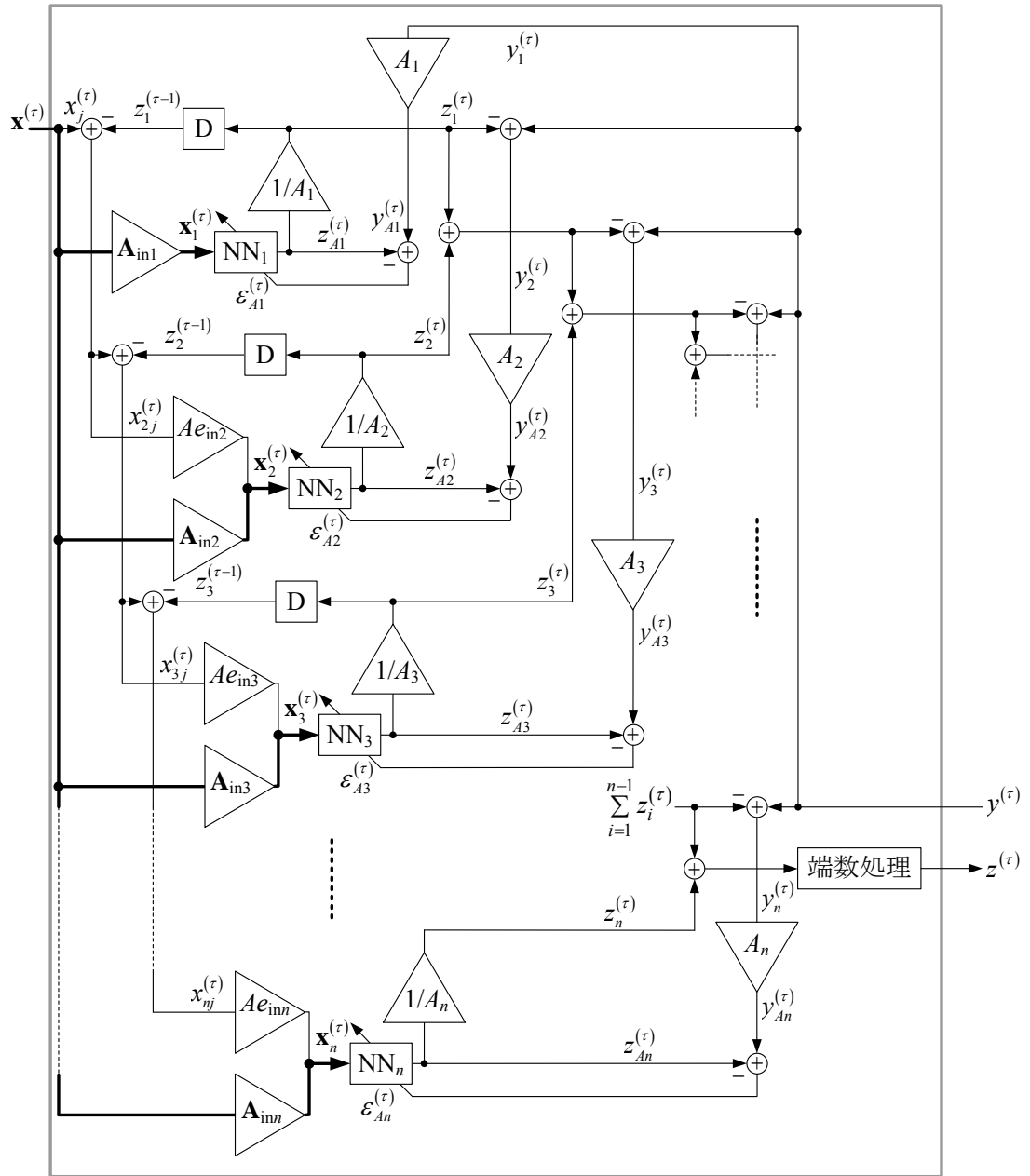


図 7.1 誤差収束型ニューロンネットワーク予測器

高精度な予測符号化を行うためには、適用原理と使用する予測器を改善しなければならない。ここで、次の理由により、予測符号化に用いられる非線形予測器としてのECNNの各段のニューロンネットワークの学習能力を改善する手段を提案する。この手段を適用した予測器をECNNPと呼ぶ。ニューロンネットワークを用いた非線形予測器の学習は、過去信号と予測信号間の相関性が強いほど容易となる。したがって、ECNNの第1段目のニューロンネットワークの学習は比較的容易である。しかし、段数が高くなるに連れ、その関数関係は弱くなると推測さ

れるため、段数が高いニューロンネットワークの学習は、難しくなる。そこで、ECNNの各段のニューロンネットワークの入力信号と教師信号の間の関数関係を強くするために、各段のニューロンネットワークを予測器として用いる。また、ニューロンネットワークの学習能力は、入力信号の増加と共に向上させることができる[3]。図7.1は、前記の改善手段を実現するために、ECNNを再設計して考案されたECNNPである。ECNNPの入出力特性を式(7.1)～(7.8)に示す。式(7.1)～(7.5)は、各ニューロンネットワークの入力を示す。式(7.4)は、第2段目以降のニューロンネットワークの初期条件である。式(7.6)は、ECNNPの各段のニューロンネットワークの入出力関係を示す。式(7.7)は、ECNNPの出力信号を示す。式(7.8)は、ECNNPの教師信号を示す。

$$n_{in} \geq 1 \quad (n_{in} \in \mathbb{Z}) \quad (7.1)$$

$$\mathbf{x}_1^{(\tau)} = (x_{A11}^{(\tau)}, x_{A12}^{(\tau)}, \dots, x_{A1n_{in}}^{(\tau)}) \quad (7.2)$$

$$\mathbf{x}_i^{(\tau)} = (Ae_{ini} x_{ij}^{(\tau)}, x_{Ai1}^{(\tau)}, x_{Ai2}^{(\tau)}, \dots, x_{Ain_{in}}^{(\tau)}) \quad (i = 2, 3, \dots, n; j = 1, 2, \dots, n_{in}) \quad (7.3)$$

$$x_{ij}^{(\tau)} = 0 \quad (\tau \leq 0; i = 2, 3, \dots, n; j = 1, 2, \dots, n_{in}) \quad (7.4)$$

$$x_{ij}^{(\tau)} = x_j^{(\tau)} - \sum_{k=1}^{i-1} z_k^{(\tau-1)} \quad \begin{pmatrix} \tau > 0 \\ i = 2, 3, \dots, n \\ j = 1, 2, \dots, n_{in} \end{pmatrix} \quad (7.5)$$

$$z_{Ai} = f_i(\mathbf{x}_i) \quad (i = 1, 2, \dots, n) \quad (7.6)$$

$$z_j^{(\tau)} = \hat{x}_j^{(\tau+1)} \quad (j = 1, 2, \dots, n_{in}) \quad (7.7)$$

$$y^{(\tau)} = x_j^{(\tau+1)} \quad (j = 1, 2, \dots, n_{in}) \quad (7.8)$$

ここで、 $x_{ij}$ はECNNPの入力信号 $x_j$ に対する第 $i$ 段目の入力誤差信号、 $Ae_{ini}$ は第 $i$ 段目の入力誤差信号の増幅率、 $f_i$ は第1段目のときは第1段目のニューロンネットワークの入出力関係を



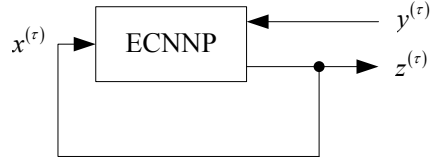


図 7.2 誤差収束型ニューロンネットワーク予測器を用いた時系列信号逐次出力器

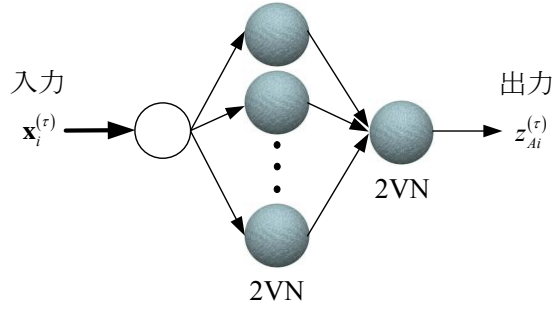


図 7.3 2 次ボルテラニューロンネットワーク

表わす  $n_{in}$  変数関数，第 2 段目以降のときは第 2 段目以降のニューロンネットワークの入出力関係を表わす  $n_{in}+1$  変数関数である．また，1 入力の ECNNP を用いた系列信号逐次出力器は，ECNNP の学習が完全ならば，実現できる．応用としては，ニューロンネットワーク付の神経振動子（Central Pattern Generator，以下 CPG と称する）として用いることができる[4]．これを図 7.2 に示す．

ECNNPに用いるニューロンネットワークは，その入力信号の特性に合わせて最適なものを選択しなければならない．ここでは，ECNNPの基本ニューロンネットワークを2VNN，または，R2VNNとする．2VNNを構成する2VNに関しては，第4章で説明している．図7.3に $n_{in}$ 入力1出力3層の2VNNを示す．また，図7.4に中間層が2VN，出力層がR2VNで構成される $n_{in}$ 入力1出力3層のR2VNNを示す．図7.5にR2VNNを構成する離散時間のR2VNを示す．R2VNの入出力特性を式(7.9)～(7.11)に示す．R2VNは2VNの1離散時間前の出力を入力へ再帰させることによって表現される．

$$u^{(\tau)} = \sum_{i=1}^n w_i^{(\tau)} x_i^{(\tau)} + w_r^{(\tau)} z^{(\tau-1)} \quad (7.9)$$

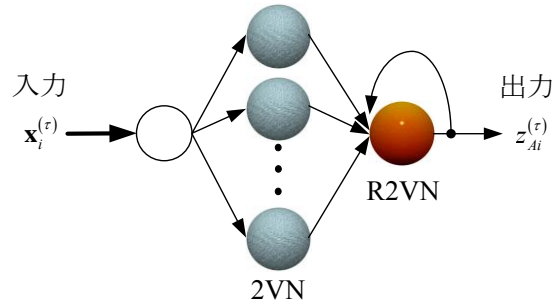


図 7.4 再帰型 2 次ボルテラニューロンネットワーク

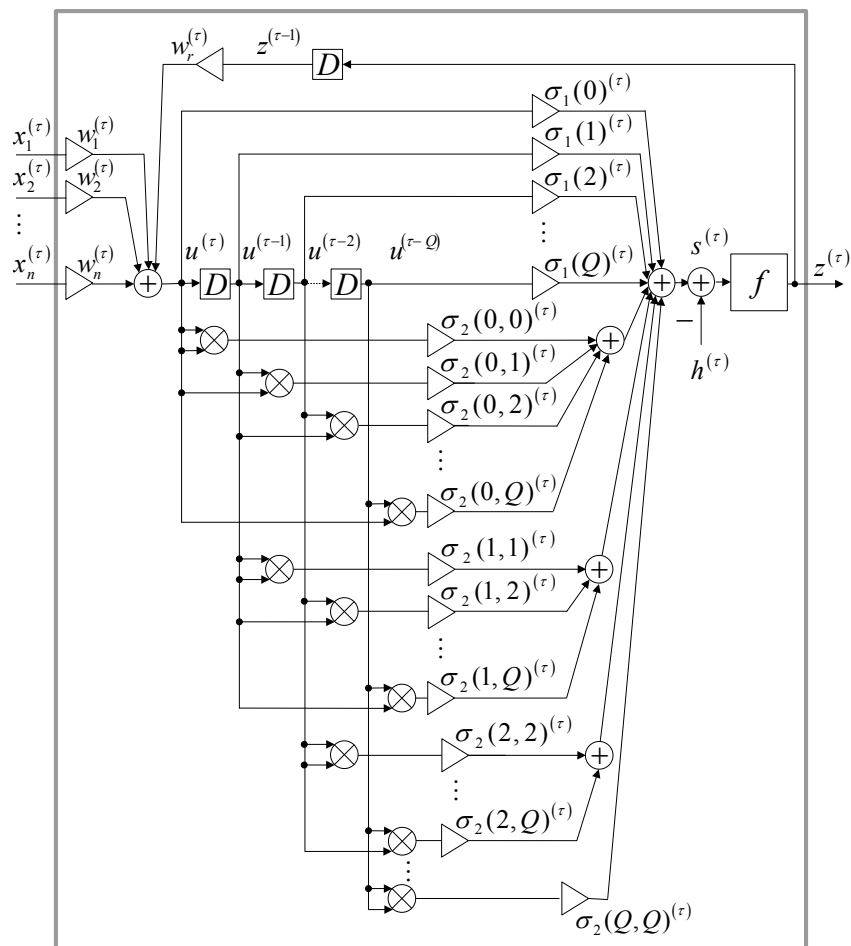


図7.5 離散時間の再帰型2次ボルテラニューロン

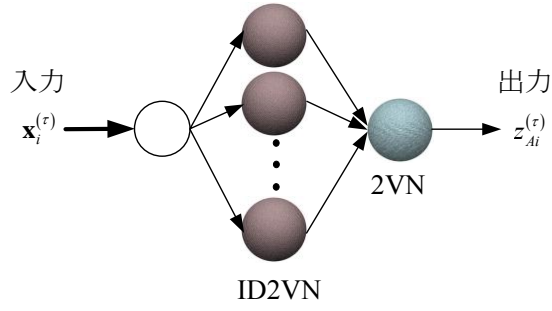


図 7.6 入力遅れ 2 次ボルテラニューロンネットワーク

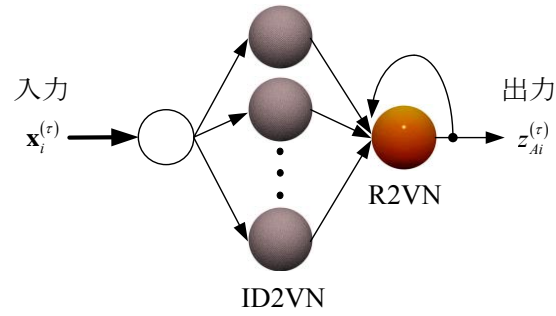


図 7.7 入力遅れ再帰型 2 次ボルテラニューロンネットワーク

$$s^{(\tau)} = \sum_{p=0}^Q \sigma_1(p)^{(\tau)} u^{(\tau-p)} + \sum_{p=0}^Q \sum_{q=p}^Q \sigma_2(p, q)^{(\tau)} u^{(\tau-p)} u^{(\tau-q)} - h^{(\tau)} \quad (7.10)$$

$$z^{(\tau)} = f(s^{(\tau)}) = A \tan^{-1}(s^{(\tau)}) \quad (7.11)$$

ここで、 $w_r$  は再帰入力のための結合荷重である． $w_i$ 、 $w_r$ 、 $h$ 、 $\sigma_1$ 、 $\sigma_2$  は学習により更新される．

基本ニューロンネットワークの能力を向上させるものとして、次の 2 つの応用ニューロンネットワークに関して述べる．図 7.6 に 2VNN の入力部に FIR フィルタを装着した ID2VNN を示す．これは、中間層が 2VN の入力部に FIR フィルタを装着した入力遅れ 2VN、以下 ID2VN と称するニューロン、出力層が 2VN で構成されている．図 7.7 に R2VNN の入力部に FIR フィルタを装着した IDR2VNN を示す．これは、中間層が ID2VN、出力層が R2VN で構成されている．

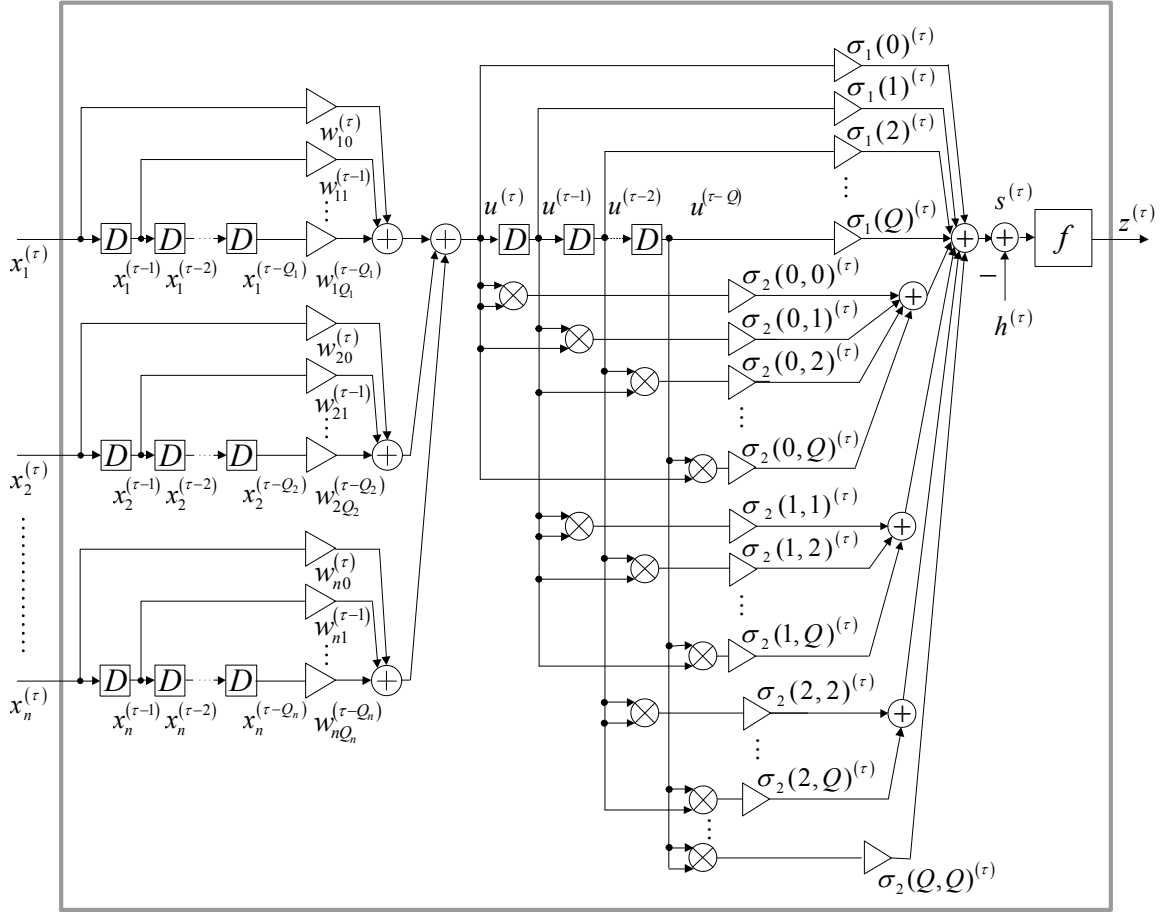


図 7.8 入力遅れ 2 次ボルテラニューロン

図 7.8 に ID2VN を示す．さらに，式(7.12)～(7.14)に ID2VN の入力特性を示す．

$$u^{(\tau)} = \sum_{i=1}^n \sum_{p=0}^{Q_i} w_{ip}^{(\tau)} x_i^{(\tau-p)} \quad (7.12)$$

$$s^{(\tau)} = \sum_{p=0}^Q \sigma_1(p)^{(\tau)} u^{(\tau-p)} + \sum_{p=0}^Q \sum_{q=p}^Q \sigma_2(p, q)^{(\tau)} u^{(\tau-p)} u^{(\tau-q)} - h^{(\tau)} \quad (7.13)$$

$$z^{(\tau)} = f(s^{(\tau)}) = A \tan^{-1}(s^{(\tau)}) \quad (7.14)$$

## 7.3 計算機シミュレーション

### 7.3.1 方法

#### a. 基本予測器の作成

計算機シミュレーションでは，図7.9に示す2段のニューロンネットワークから構成される ECNNPへ空間方向が1次元の時系列パターンを入力信号 $x(\tau)$ と教師信号 $y(\tau)=x(\tau+1)$ の組合せを与

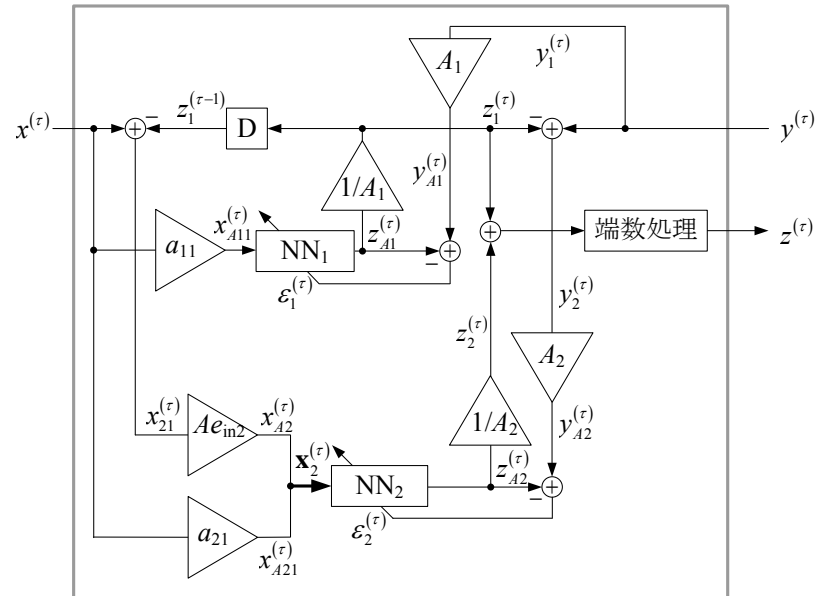


図 7.9 2 段のニューロンネットワークで構成されている誤差収束型ニューロンネットワーク 予測器

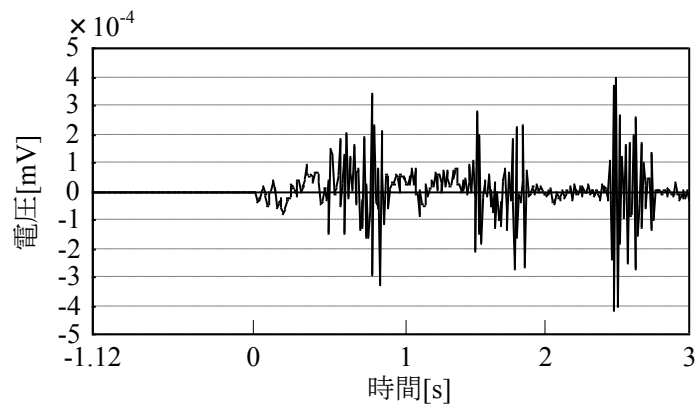


図 7.10 第 2 段目のニューロンネットワークの学習用信号

表7.1 第2段目のニューロンネットワークの学習計算機シミュレーション条件

学習則				2VNNの学習則	
初期値		結合荷重		-0.3～0.3	
		しきい値		-0.3～0.3	
		予測係数	$\sigma_1$	$0.7 \times 0.3^p$	
$\sigma_2$	$0.7 \times 0.3^p \times 0.7 \times 0.3^q$				
中間層素子数			範囲	2～20	
			間隔	2	
タップ数, または フィルタ長			範囲	2～56	
			間隔	2	
学習 強化 係数	勾配法項	粗探索	範囲	$10^{-4} \sim 1$	
			間隔	10倍	
		細探索	範囲	$0.2 \alpha_{min} \sim 0.9 \alpha_{min}, 2 \alpha_{min} \sim 9 \alpha_{min}$	
			間隔	$0.1 \alpha_{min}, \alpha_{min}$	
	慣性項			0	
	出力係数			1	
学習サイクル			30,000		
施行回数			3		

え学習させる．ここで，ECNNPの学習は，1段ごと学習を独立して仕上げる逐次学習とする．

はじめに，ECNNPのニューロンネットワークとして2VNNを用いる．第1段目のニューロンネットワーク，以下 $NN_1$ と称するニューロンネットワークへ空間方向が1次元の時系列パターンの入力信号 $x(\tau)$ と教師信号 $y_1(\tau)=x(\tau+1)$ の組合せを与え学習させる．ここで， $A_{in1}=1$ ， $A_1=1$ とし，学習には，第4章の図4.6の健常者の安静時における正常洞調律心電図信号を用いた図4.5に示す1入力1出力3層の2VNNの学習結果を用いる．この細探索学習で，最小平均RMSEを実現した中間層素子数とフィルタ長を用いた2VNNを $NN_1$ とする．さらに，この $NN_1$ の出力信号とECNNPの教師信号から得られた誤差信号を第2段目のニューロンネットワーク，以下 $NN_2$ と称するニューロンネットワークの学習用信号の一部分とする．この学習用信号を図7.10に示す．

次に， $NN_2$ へ空間方向が2次元の時系列パターンの入力信号 $x(\tau)$ ， $x_{21}(\tau)$ と空間方向が1次元の時系列パターンの教師信号 $y_2(\tau)=x_{21}(\tau+1)$ の組合せを与え学習させる．入力信号 $x(\tau)$ には，図4.6の心電図信号，入力信号 $x_2(\tau)$ と教師信号 $y_2(\tau)$ には，式(7.3)の初期条件と図7.10の学習用信号から得られる信号を用いる．ここで， $Ae_{in2}=A_2$ ， $a_{21}=1$ とし， $A_2$ には，図7.10の学習用信号の0秒から始まる誤差信号部分の最大絶対値を1にする値を用いる．学習は，112離散時間のデータを初期入力後，入力信号と教師信号の対を1回提示する．この過程を1学習サイクルとし，これを繰り返す．計算機シミュレーションは，はじめに，中間層素子数とフィルタ長を計算機シミュレーション条件設定パラメータとし，それらに対して，勾配法項の学習強化係数を変化させなが

ら、3回の粗探索学習を行い、その結果得られた平均RMSEを比較する。次に、最小平均RMSEを実現した中間層素子数を用いて、フィルタ長を計算機シミュレーション条件設定パラメータとし、それらに対して、勾配法項の学習強化係数を、 $\alpha_{min}$ を元に変化させながら、3回の細探索学習を行う。 $\alpha_{min}$ は、粗探索学習で中間層素子数とフィルタ長の組合せごとに、最小平均RMSEを得たときの勾配法項の学習強化係数である。その結果得られた平均RMSEを比較する。そこで、最小平均RMSEを実現した中間層素子数とフィルタ長を用いた2VNNをNN<sub>2</sub>とする。そして、NN<sub>2</sub>の出力信号をECNNPへ与えた教師信号のレベルへ復元し、その信号と同時刻のNN<sub>1</sub>の出力信号に加算する。最後に、この加算後の出力信号の小数点以下5桁目を四捨五入する端数処理を行う。その結果得られる出力信号をECNNPの出力信号とし、学習精度を評価する。

表7.1にNN<sub>2</sub>の学習計算機シミュレーション条件を示す。ここで、予測係数の初期値は、指数平滑法を用いる。その他の初期値は、学習計算機シミュレーション1回ごとに、擬似乱数で決定される。

#### b. 基本予測器の学習汎用性の評価

基本予測器の学習汎用性を評価するために、先の学習計算機シミュレーションで得られた基本予測器に用いた各段の2VNNの諸元を基に構成した2VNNを用いた別のECNNPと正常洞調律心電図を用いて、学習計算機シミュレーションを行う。学習用信号として、MIT-BIH No.16786の正常洞調律心電図信号の開始から940個、または、2,220個を用いる。この心電図信号は、標

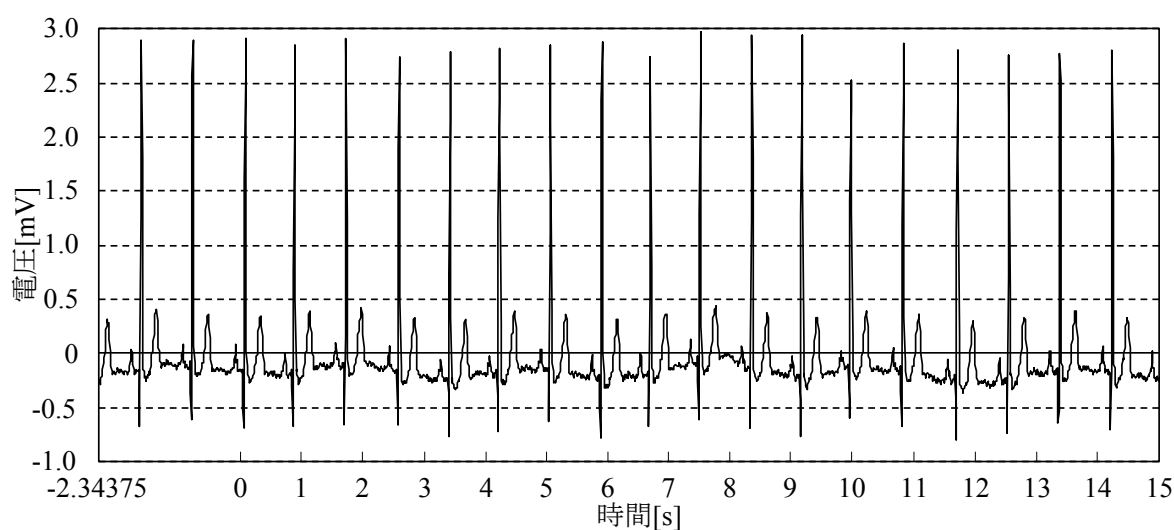


図7.11 学習汎用性確認用学習用心電図信号

表7.2 学習汎用性と学習能力改善の評価のための学習計算機シミュレーション条件

種 類			2VNN		ID2VNN	
段 数			1段目	2段目	1段目	2段目
学習則			2VNNの学習則			
初期値	結合荷重		-0.3~0.3			
	しきい値		-0.3~0.3			
	予測	$\sigma_1$	$0.7 \times 0.3^p$			
	係数	$\sigma_2$	$0.7 \times 0.3^p \times 0.7 \times 0.3^q$			
中間層素子数			4	10	4	10
フィルタ長			69	64	69	64
タップ数			1		142	
学習	勾配	範囲	$10^{-5} \sim 1$			
強化	法項	間隔	10倍			
係数	慣性項		0			
出力係数			1			
心電図信号長			5, 15秒		15秒	
学習サイクル			30,000			
試行回数			3			

本周波数128Hz, 量子化幅 $5 \times 10^{-3}$  mVで記録されている. この心電図信号を図7.11に示す. ECNNPのニューロンネットワークの構成は, 基本予測器のニューロンネットワークの緒元を基に, 中間層数, 素子数, フィルタ長を決定する. フィルタ長は, 標本周波数の倍率に従い, 基本予測器のニューロンネットワークのフィルタ長の1.28倍とする.

はじめに, 5秒間の正常洞調律心電図信号を用いた学習計算機シミュレーションを行う. 次に, さらに信号データ長を長くした学習計算機シミュレーションとして, 15秒間の正常洞調律心電図信号を用いた学習計算機シミュレーションを行う. 先の計算機シミュレーションで用いた心電図信号とここでの心電図信号の異なる点は, 被験者, 標本周波数, 量子化幅, 信号データ長, 最大振幅である. また,  $A_{in1}=A_1=1/3.275$ ,  $Ae_{in2}=A_2$ ,  $a_{21}=1/3.275$ とし,  $Ae_{in2}$ ,  $A_2$ には,  $NN_2$ の学習用誤差信号の最大絶対値を1にする値を用いる. 学習は, 300離散時間のデータを初期入力後行われる. この学習条件を表7.2に示す.

### c. 学習能力の改善

2VNNを用いたECNNPの学習能力を改善するために, ID2VNNを用いたECNNPと正常洞調律心電図を用いて, 学習計算機シミュレーションを行う. 学習用信号は, 学習汎用性を確認するときと同じ図7.11のMIT-BIH No.16786の正常洞調律心電図信号の開始から2,220個を用いる. ここでのECNNPのニューロンネットワークの構成も, 前記と同様に基本予測器のニューロンネットワークの緒元を基にし, 中間層数, 素子数, フィルタ長を決定する. ただし, タップ数は, 標



本周波数の倍率に従い、第4章で得られた最良のIDNNのタップ数の1.28倍とする。そして、15秒間の正常洞調律心電図信号を用いた学習計算機シミュレーションを行う。また、 $A_{in1}=A_1=1/3.275$ ,  $A_{e_{in2}}=A_2$ ,  $a_{21}=1/3.275$ とし、 $A_{e_{in2}}$ ,  $A_2$ には、 $NN_2$ の学習用誤差信号の最大絶対値を1にする値を用いる。学習は、300離散時間のデータを初期入力後行われる。この学習条件を表7.2に示す。

### 7.3.2 結果

#### a. 基本予測器の作成

はじめに、基本予測器の学習計算機シミュレーションの結果を示す。図 7.12 と 7.13 は、 $NN_2$ の学習計算機シミュレーション後に得られた平均 RMSE とその標準偏差を示す。ここで、図中の縦棒は、標準偏差の $\pm$ の範囲を示す。図 7.12 より、粗探索学習では、中間層素子数が少なく、フィルタ長が長い部分に、平均 RMSE が小さくなる良好な学習結果が集まっていることが確認できる。細探索学習では、図 7.13 に示すように学習強化係数が  $10^{-2}$  のときの 1 箇所著しく平均 RMSE が減少する。図 7.14 は、 $NN_2$  の 3 回の細探索学習で、最小平均 RMSE が得られたときの学習開始前、学習サイクル 1 回目、学習サイクル 100 回ごとに記録した学習サイクルと 1 単位時間データ当りの平均評価関数値の平均との関係を示す。この図より、学習サイクルが 25,000 回付近までは、1 単位時間データ当りの平均評価関数値の平均は指数関数的に減少しており、学習が良好に行われていることが確認できる。しかし、学習サイクルが 25,000 回付近を越えると、1 単位時間データ当りの平均評価関数値の平均は横ばいとなる。これより、 $NN_2$  の学習が飽和する学習サイクルが確認できる。図 7.15 に、細探索学習で最小平均 RMSE が得られたときの平均出力信号と教師信号を比較した結果を示す。この図の教師信号と平均出力信号との間には誤差があるものの、視覚的には完全に一致しており、学習の良好な仕上がりで安定した高精度な出力が得られていることが確認できる。さらに、図 7.16 に、このときの平均予測誤差系列をそれぞれ示す。予測誤差系列は、教師信号値から平均出力信号値を引いて得た差の系列である。この図より、低周波のうねりが確認できる。しかし、その誤差が ECNNP の教師信号の有効数字の桁数より非常に小さいため、精度上の問題にはならない。ここで、ネットワークの条件は、中間層素子数 8、フィルタ長 50、学習強化係数  $1 \times 10^{-2}$  である。図 7.17 に  $NN_1$  と  $NN_2$  の出力信号を示す。図 7.18 に  $NN_1$  と  $NN_2$  の出力信号を加算し、端数処理を行った結果得られた ECNNP の出力信号を示す。この図の教師信号と出力信号は、完全に一致しており、全

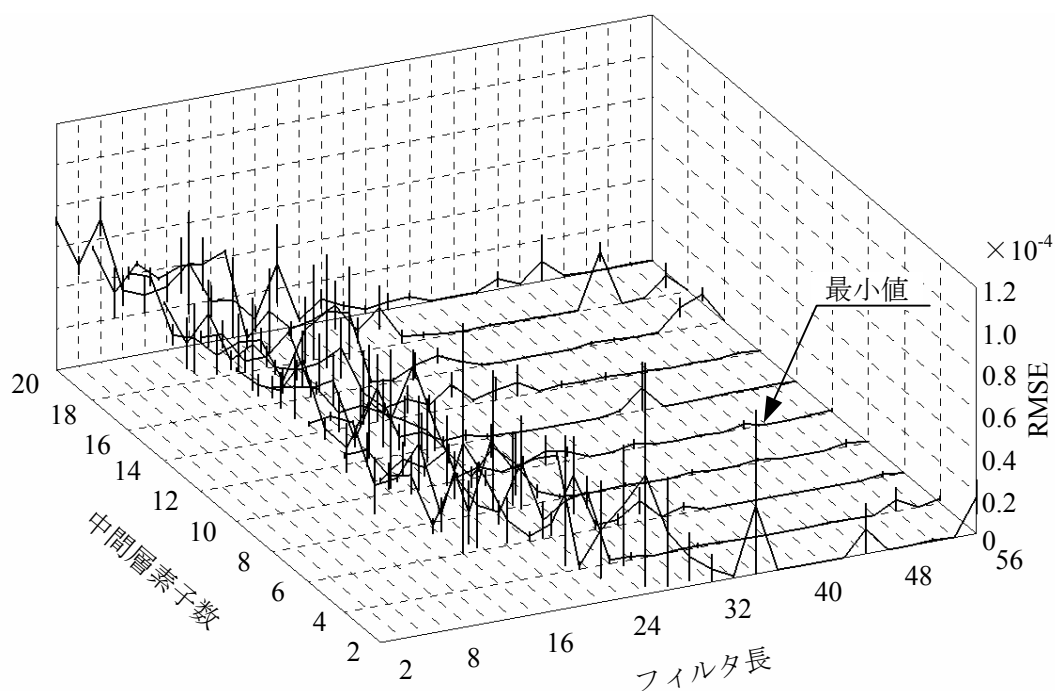


図 7.12 第 2 段目のニューロンネットワークの粗探索学習後の平均 RMSE と標準偏差

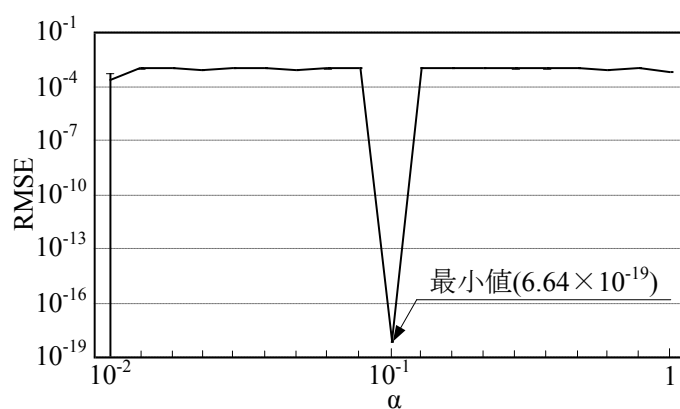


図7.13 第2段目のニューロンネットワークの細探索学習後の勾配法項学習強化係数 $\alpha$ に対する平均RMSEと標準偏差

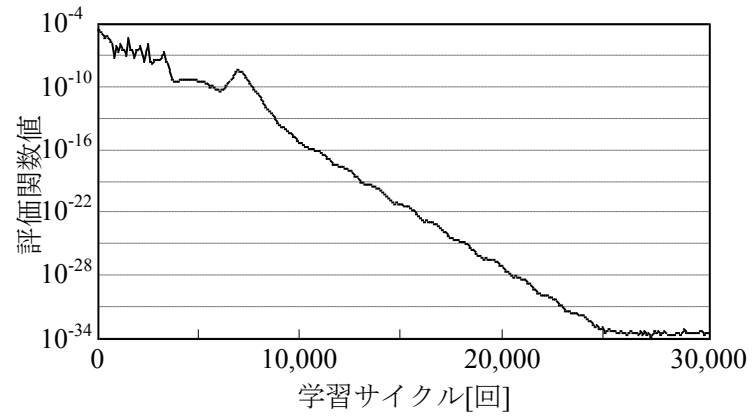


図7.14 第2段目のニューロンネットワークの最小平均RMSEのときの学習サイクルと1単位時間データ当りの平均評価関数値の平均との関係

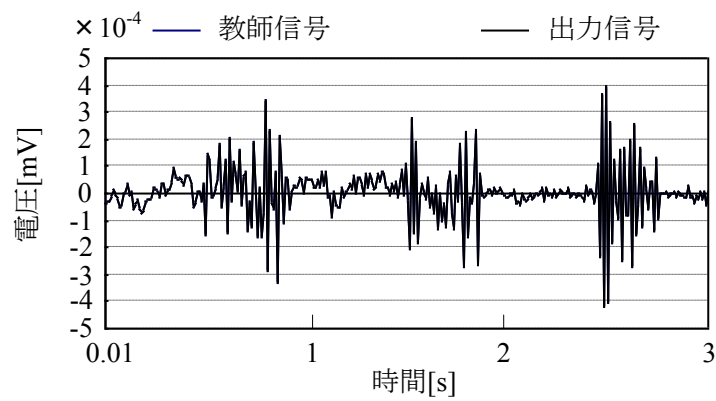


図7.15 第2段目のニューロンネットワークの最小平均RMSEのときの平均出力信号

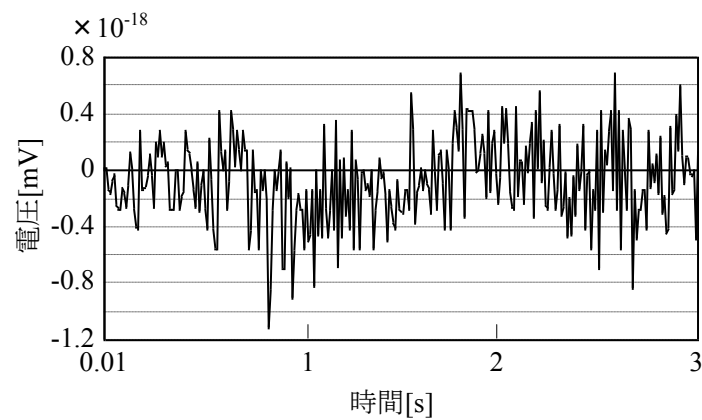


図7.16 第2段目のニューロンネットワークの最小平均RMSEのときの平均誤差系列

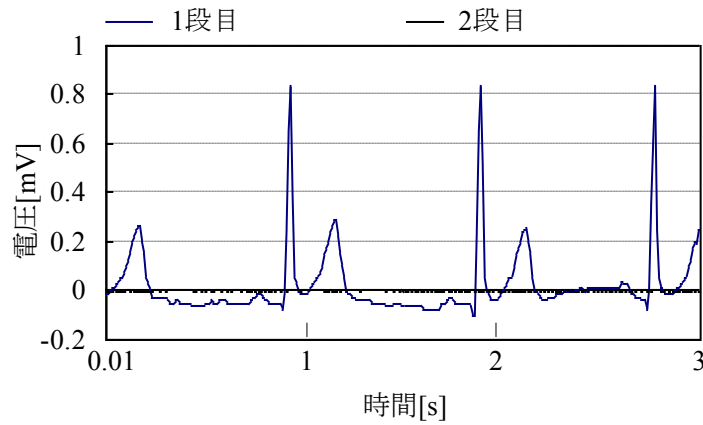


図7.17 第1段目と第2段目のニューロンネットワークの出力信号

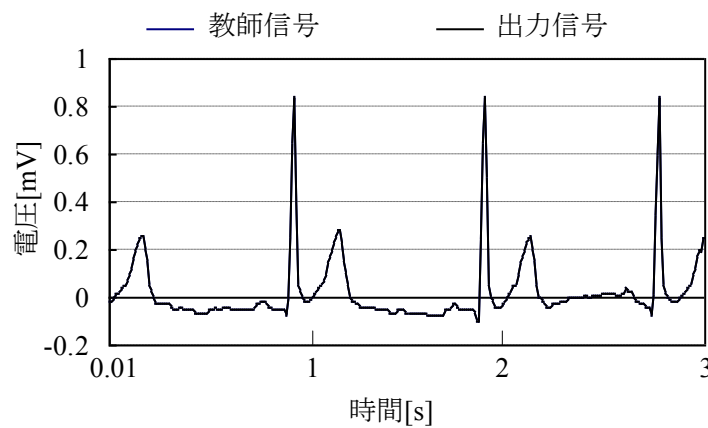


図7.18 誤差収束型ニューロンネットワーク予測器の出力信号

く誤差のない出力信号が得られている。以上より、ECNNP の完全学習能力を実証できた。そして、ECNNP の汎化能力の検証はないものの、ECNNP が高精度予測器になりえる実現性はあるといえる。さらに、この実証により、図 7.2 に示す時系列信号逐次出力器の実証もできた。

#### b. 基本予測器の学習汎用性の評価

はじめに、5 秒間の正常洞調律心電図信号を用いて行った 2VNN で構成される ECNNP の学習計算機シミュレーションの結果を示す。図 7.19 は、 $NN_1$  の学習後、最小平均 RMSE を実現したときの出力信号を基に得た  $NN_2$  の学習用誤差信号である。図 7.20 より、最小平均 RMSE を実現したときの各段のニューロンネットワークの勾配法項の学習強化係数が得られる。図 7.21 は、各ニューロンネットワークの 3 回の学習で、最小平均 RMSE が得られたときの学習開始前、学習サイクル 1 回目、学習サイクル 100 回ごとに記録した学習サイクルと 1 単位時間デ

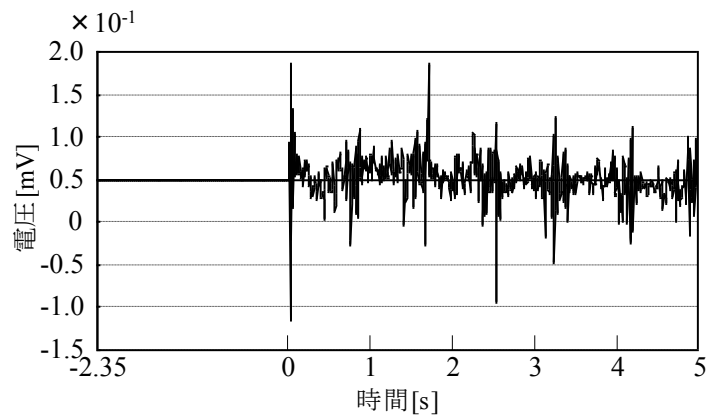


図 7.19 第 2 段目のニューロンネットワーク学習用誤差信号

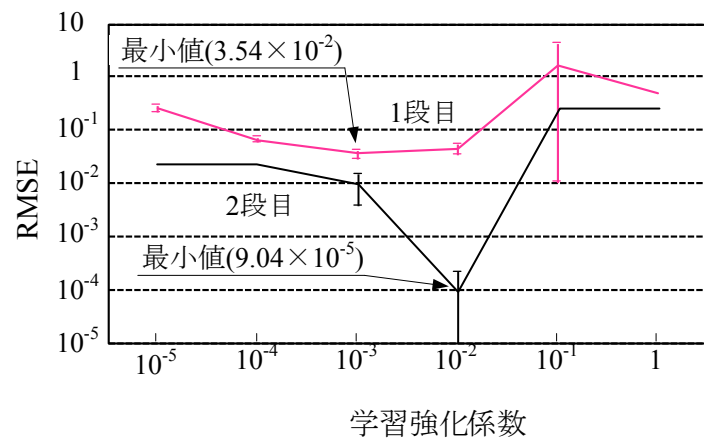


図 7.20 勾配法項学習強化係数に対する平均 RMSE と標準偏差

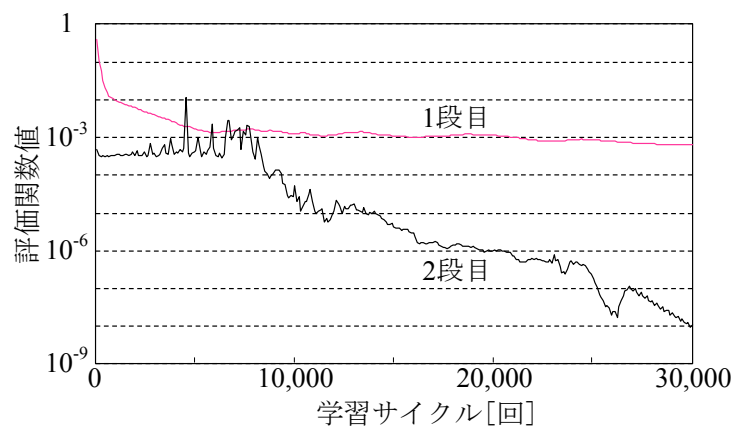


図 7.21 学習サイクルと 1 単位時間データ当りの平均評価関数値の平均との関係

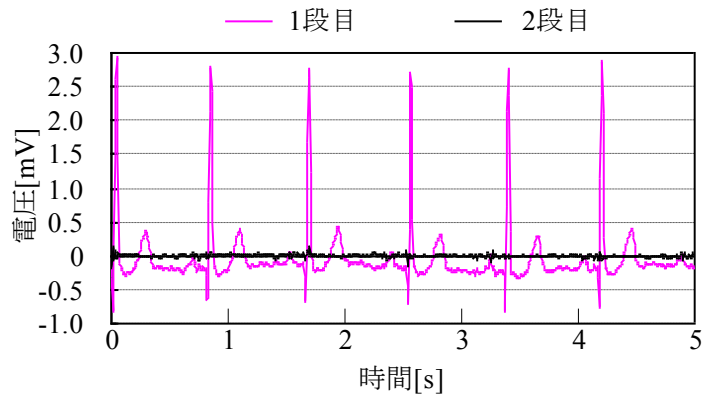


図 7.22 第1段目と第2段目のニューロンネットワークの出力信号

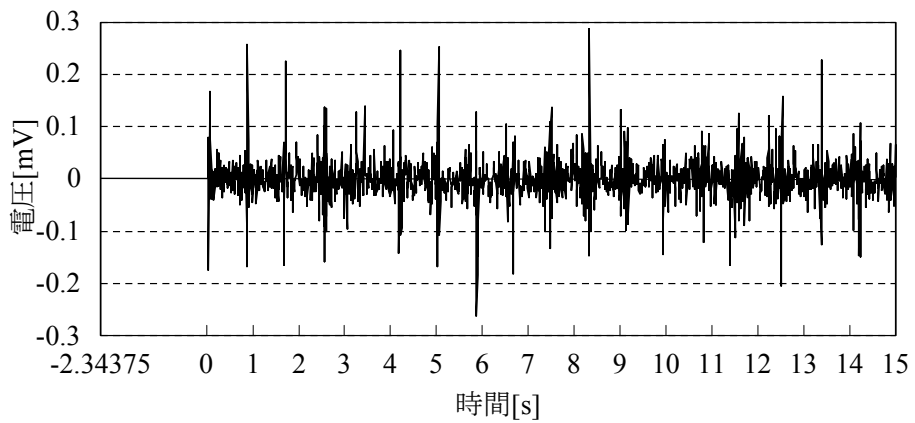


図 7.23 第2段目のニューロンネットワーク学習用誤差信号

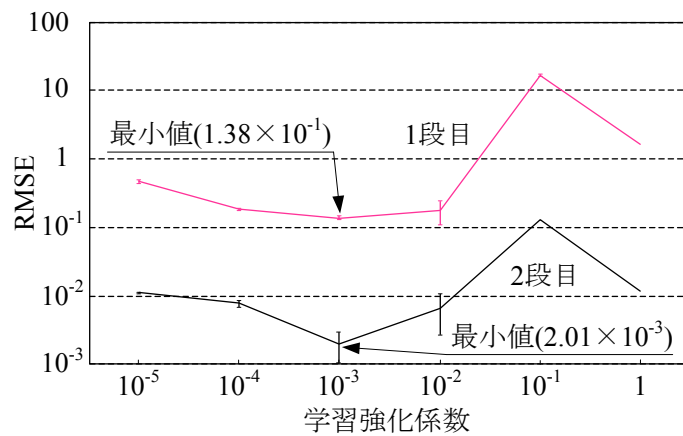


図 7.24 勾配法項学習強化係数に対する平均 RMSE と標準偏差

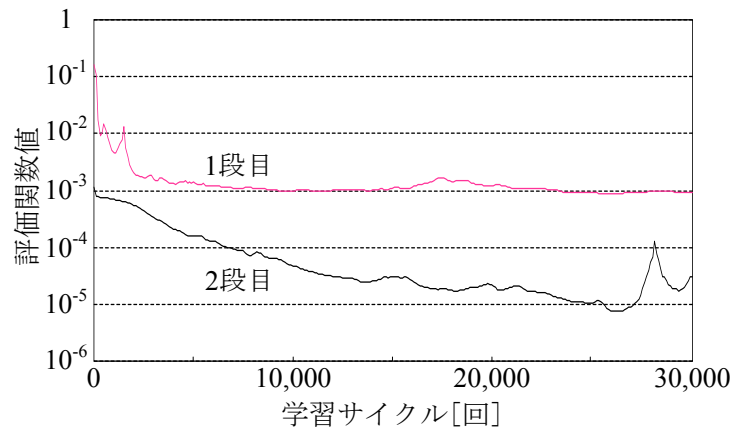


図 7.25 学習サイクルと 1 単位時間データ当りの平均評価関数値の平均との関係

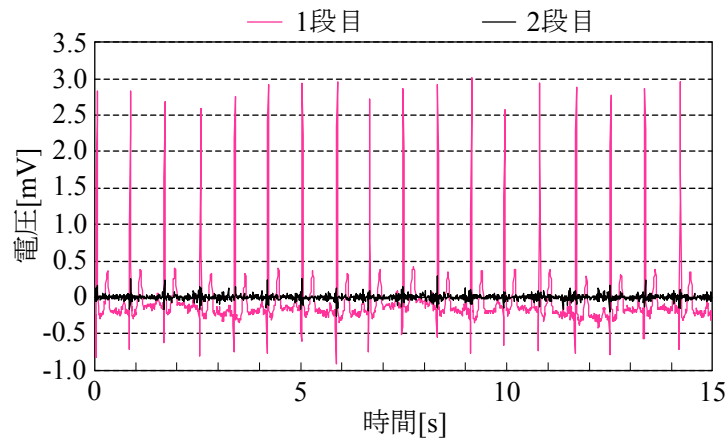


図 7.26 第 1 段目と第 2 段目のニューロンネットワークの出力信号

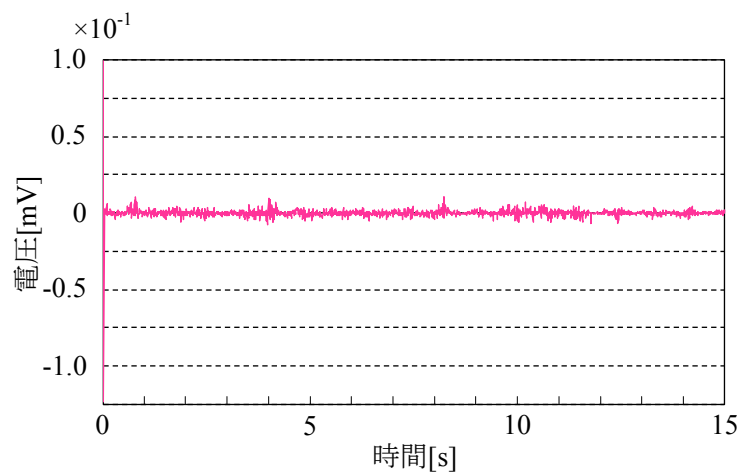


図 7.27 誤差収束型ニューロンネットワーク予測器の誤差信号

一タ当りの平均評価関数値の平均との関係を示す。この図は、 $NN_1$  の学習が飽和したときの予測誤差を、 $NN_2$  を用いることで、さらに小さくすることができることを示している。また、学習サイクルが 3 万回での  $NN_2$  の 1 単位時間データ当りの平均評価関数値の平均の勾配は、さらに学習が可能であるかのように推測できる。図 7.22 に  $NN_1$  と  $NN_2$  の出力信号を示す。これらの信号の和を端数処理することで、誤差がない ECNNP の教師信号と同一の出力信号が得られる。以上より、ECNNP の予測誤差がない完全学習が得られ、正常洞調律心電図信号に対する良好な学習能力の確認ができた。

次に、15 秒間の正常洞調律心電図信号を用いて行った 2VNN で構成される ECNNP の学習計算機シミュレーションの結果を示す。図 7.23 は、 $NN_1$  の学習後、最小平均 RMSE を実現したときの出力信号を基に得た  $NN_2$  の学習用誤差信号である。図 7.24 より、最小平均 RMSE を実現したときの各段のニューロンネットワークの勾配法項の学習強化係数が得られる。図 7.25 は、各ニューロンネットワークの 3 回の学習で、最小平均 RMSE が得られたときの学習開始前、学習サイクル 1 回目、学習サイクル 100 回ごとに記録した学習サイクルと 1 単位時間データ当りの平均評価関数値の平均との関係を示す。この図は、 $NN_1$  の学習が飽和したときの予測誤差を、 $NN_2$  を用いることで、さらに小さくすることができることを示している。しかし、図 7.21 の 5 秒間の正常洞調律心電図信号を用いた結果より悪く、学習が教師信号の長さに従って、より困難になっていることが確認できる。また、学習サイクルが 3 万回での  $NN_2$  の 1 単位時間データ当りの平均評価関数値の平均の勾配は、飽和している。図 7.26 に  $NN_1$  と  $NN_2$  の出力信号を示す。これらの信号の和を端数処理しても、誤差がない教師信号は得られない。このため、図 7.27 に示す ECNNP の予測誤差が残る。この誤差信号をなくすためには、ニューロンネットワークの段数を高くする必要がある。

### c. 学習能力の改善

15 秒間の正常洞調律心電図信号を用いて行った ID2VNN で構成される ECNNP の学習計算機シミュレーションの結果を示す。図 7.28 は、 $NN_1$  の学習後、最小平均 RMSE を実現したときの出力信号を基に得た  $NN_2$  の学習用誤差信号である。図 7.29 より、最小平均 RMSE を実現したときの各段のニューロンネットワークの勾配法項の学習強化係数が得られる。図 7.30 は、各ニューロンネットワークの 3 回の学習で、最小平均 RMSE が得られたときの学習開始前、学習サイクル 1 回目、学習サイクル 100 回ごとに記録した学習サイクルと 1 単位時間データ当りの



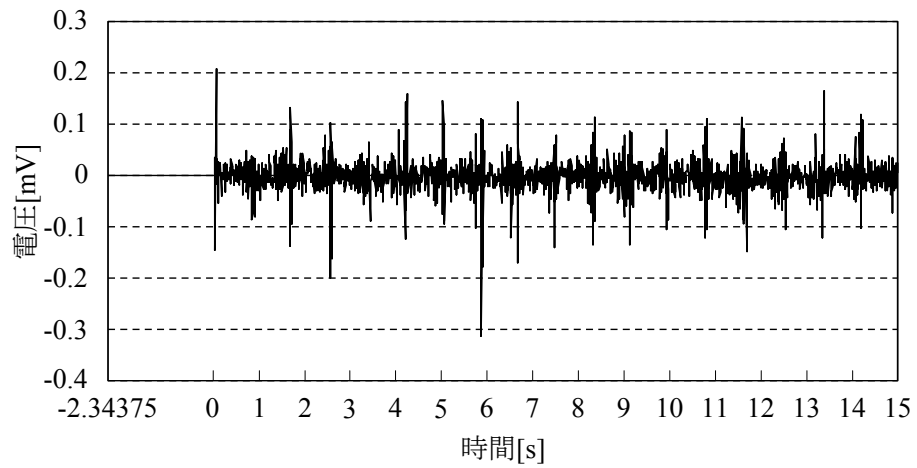


図 7.28 第 2 段目のニューロンネットワーク学習用誤差信号

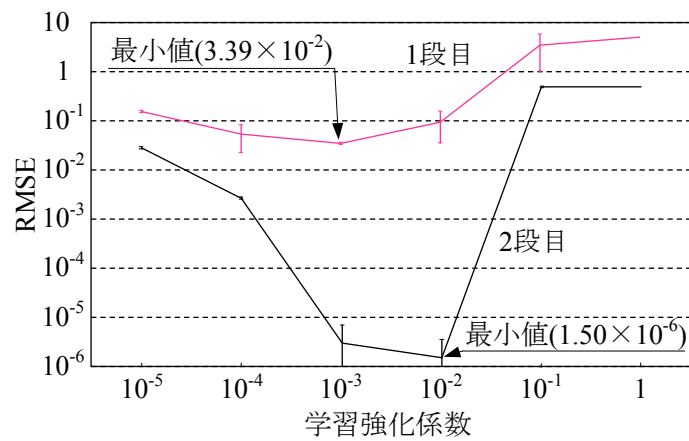


図 7.29 勾配法項学習強化係数に対する平均 RMSE と標準偏差

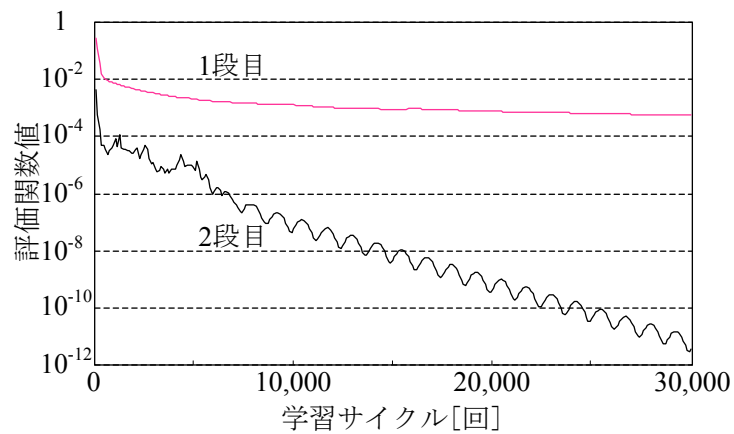


図 7.30 学習サイクルと 1 単位時間データ当りの平均評価関数値の平均との関係

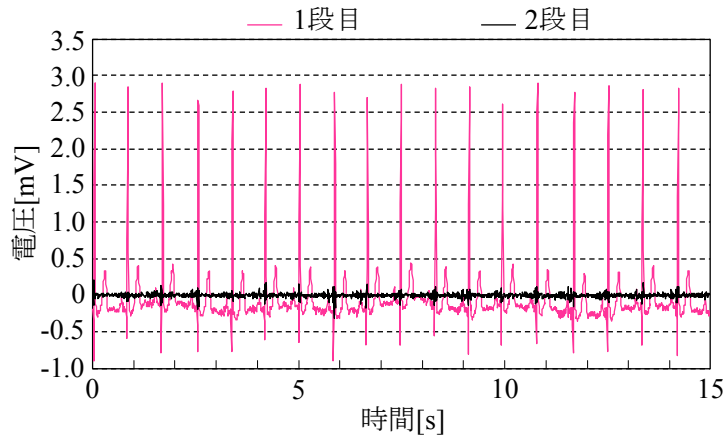


図 7.31 第 1 段目と第 2 段目のニューロンネットワークの出力信号

平均評価関数値の平均との関係を示す．この図は， $NN_1$  の学習が飽和したときの予測誤差を， $NN_2$  を用いることで，さらに小さくすることができることを示している．また，学習サイクルが 3 万回での  $NN_2$  の 1 単位時間データ当りの平均評価関数値の平均の勾配は，さらに学習が可能であるかのように推測できる．図 7.31 に  $NN_1$  と  $NN_2$  の出力信号を示す．これらの信号の和を端数処理することで，誤差がない ECNNP の教師信号と同一の出力信号が得られる．このように ID2VNN を ECNNP に用いることで，完全学習が成し遂げられ，図 7.27 に示したような ECNNP の予測誤差は生じなくなる．以上より，2VNN で構成される ECNNP の正常洞調律心電図信号に対する学習能力を ID2VNN で構成される ECNNP で改善できることが示された．

## 7.4 考察

はじめに，ECNNP の学習能力について考察する．ECNNP は，入力信号と教師信号の関数関係が弱くなり，学習が困難となる第 2 段目以降のニューロンネットワークの学習を可能とするために，入力信号と出力信号の関数関係を利用できるようにしたニューロンネットワークを ECNNP の内部予測器として用いることで，学習能力を改善したものである．さらに，入力信号に対する出力信号の一意性を強化するために，2 つの入力信号を用いている．このため，2 段の 2VNN で構成される ECNNP を 3 秒間の正常洞調律心電図信号を用いて学習した結果，完全学習が可能であった．

次に，基本予測器の学習汎用性について考察する．正常洞調律心電図信号は，被験者が異な

っても、類似性が高く、一つの時系列信号パターンとしてとらえることができる。このため、前記の完全学習済みの ECNNP、すなわち基本予測器に用いた各段の 2VNN の諸元を基に構成した 2VNN を用いた別の ECNNP でも、ある程度完全学習は可能であると予想されていた。実際には、異なる条件の 5 秒間の正常洞調律心電図信号では、完全学習が可能であったが、15 秒間のものでは、不可能であった。これは、学習で用いた心電図信号長が基本予測器で用いた信号長の 1.5 倍程度では問題はないが、3 倍程度となると学習能力に問題が生じるということである。学習で用いる信号長が、学習済みの信号長より長くなるに伴って、基本予測器の諸元を基本としたものでは、能力不足になっている。しかし、ID2VNN を ECNNP に用いることで、15 秒間の正常洞調律心電図信号の完全学習は可能となった。この理由は、ID2VNN の方が 2VNN より、ネットワークの入力信号数やモデルパラメータ数が多く、学習能力が高いためであると考えられる。したがって、ID2VNN を用いた ECNNP は、正常洞調律心電図信号に対して学習能力に優れていると考えられる。

今回の計算機シミュレーションは、ECNNP の完全学習遂行能力を実証したが、未学習信号に対する汎化能力を確認しておらず、高精度予測器の実証としては不十分である。このため、不整脈の心電図等を用いた汎化能力の検証が必要である。

最後に、学習能力の改善について考察する。ECNNP の各段のニューロンネットワークの入力信号と教師信号の間に関数関係がほとんどない場合、ECNNP の入出力信号の関係の学習は、難しくなる。さらに、入出力信号に確定信号と不規則信号の成分が含まれていると推測される。したがって、学習を改善する手段として、入出力信号のこれらの 2 成分についてそれぞれ考えることが必要である。

確定信号成分に対する改善としては、ニューロンネットワークの出力を再帰させる手段がある。これは、ニューロンネットワークの過去の出力信号から現在の出力信号を自己生成する機能、すなわち、自己相関をニューロンネットワークへ与えることで学習能力を改善する。R2VNN は、この手段を適用するために、出力層に R2VN を用いて、出力信号が出力層の入力として再帰できるようにしたものである。また、ネットワークの入力信号数やモデルパラメータ数を多くすることができる入力遅れや時間遅れの構造を付加することで、学習能力をさらに改善することができると考えられる。これらの改善により、長時間の心電図信号に関して、完全学習が得られるならば、ECNNP だけで、その情報の保存機能を実現できる可能性がある。また、学習能力が高いと考えられる前記のニューロンネットワークを用いた ECNNP1 台だけで、多数の

被験者の心電図信号を多重保存する機能を実現できる可能性もある。

不規則信号成分に対する改善としては、不規則信号を雑音として取り扱う方法がある。これは、前処理として ECNNP の入力信号と教師信号をフィルタリングすることで、これらの雑音を除去する。この方法に適用できるフィルタリングとしては、フーリエ変換、ベイズモデル、トレンドモデル等[5]がある。しかし、不規則信号の中でも、自然、生体、経済等の信号情報のように必要なものもある。このため、この方法は、必要な情報を欠落させないように注意して行わなければならない。

## 7.5 結言

本章では、ECNN の各段のニューロンネットワークの入力信号と教師信号の間の関数関係がほとんどない場合の学習困難性を、ECNN の各段のニューロンネットワークを予測器として構成し、その入力信号と教師信号の間の相関性を強くすることで改善した ECNNP を考案した。そして、ECNNP の各段のニューロンネットワークへ 2VNN, R2VNN, ID2VNN, IDR2VNN を用いることを提案した。

次に、正常洞調律心電図信号を用いて、2 段の 2VNN から構成される ECNNP の学習計算機シミュレーションを行い、その学習精度を評価した。その結果、誤差のない出力を得る学習能力、すなわち、完全学習能力が ECNNP にあることを実証した。この実証により、ECNNP を用いた時系列信号逐次出力器の実現も示された。これは、ニューロンネットワーク付き CPG としても応用することができる。また、ECNNP を高精度予測器として実現できる可能性があることも示された。さらに、この ECNNP を基本予測器とし、その緒元を基に構成した別の 2 段の 2VNN で構成される ECNNP と条件の異なる正常洞調律心電図信号を用いて学習能力を評価した。その結果、学習で用いた心電図信号長が、基本予測器で用いた信号長の 1.5 倍程度では、問題なく完全学習が成し遂げられたが、3 倍程度となると完全学習ができず、学習能力に問題が生じた。この学習能力を改善するために、2 段の ID2VNN で構成される ECNNP と完全学習ができなかった心電図信号を用いて、学習計算機シミュレーションを行った結果、完全学習を成し遂げることができた。これにより、2VNN で構成される ECNNP の正常洞調律心電図信号に対する学習能力を ID2VNN で構成される ECNNP で改善できることが示された。

しかし、ECNNP は、未学習信号に対する汎化能力が確認されていないため、高精度予測器

としての十分な実証ができていない．今後の課題は，不整脈心電図等を用いて，ECNNP の汎化能力の評価を行うことである．

## 参考文献

- [1] 小早川俊祐, 藤井崇文, 横井博一, “心電図に対するニューロンネットワークの非線形予測能力の評価,” BMFSA 第 20 回年次大会講演論文集, pp.9-12, 2007.
- [2] A. Chatterjee, A. Nait-Ali and P. Siarry, “An Input-delay Neural-Network-Based Approach for Piecewise ECG Signal Compression,” IEEE Transactions on Biomedical Engineering, vol. 52, no. 5, pp. 945-947, 2005.
- [3] 小早川俊祐, 横井博一, “入力信号数に対する BP ネットワークの学習能力の評価,” 信学技報, SANE2007-102-124, vol.107, no.442, pp.83-86, 2008.
- [4] S. Kobayakawa and H. Yokoi, “Proposal of Error Convergence-type Predictor Using 2nd-order Volterra Neuron Networks with Input-delay,” Abstracts of International Symposium on Application of Biomechanical Control Systems to Precision Engineering, Fukushima, Japan, pp.114-121, 2010.
- [5] 小野寺光永, 井須芳美, 長嶋雲兵, 吉田裕亮, 細矢治夫, 永川祐三, “フーリエ変換とベイズモデルおよびトレンドモデルによる時系列データのノイズ除去,” The Journal of Chemical Software, vol.5, no.3, pp.113-127, 1999.

## 第8章 誤差収束型ニューロンネットワーク予測器を用いた予測符号化

### 8.1 緒言

ECNNPは、応用として、予測符号化の予測器に用いることができる[1]。予測符号化は、予測器の汎化精度が高いほどデータ圧縮率が向上し、その性能評価は高いものとなる。ECNNPをこの予測器として用いる場合、ECNNPの汎化精度、すなわち、汎化能力に関して評価されていなければならない。前章で、ECNNPは特定の条件の下で、完全学習ができることが実証されている。しかし、その汎化能力に関しては、評価されていない。

ECNNPの学習方法は、大きく分けるとオンライン学習とオフライン学習がある。各段のニューロンネットワークに関する学習は、同時刻のデータ1個ごと各段のニューロンネットワークで学習を行う同時学習と1段ごと学習を独立して仕上げる逐次学習である。ここでは、逐次学習後に得られる汎化能力の応用に関して述べる。前章のECNNPの学習方法は、単純に前段のニューロンネットワークの教師信号と出力信号から得られる誤差信号を後段のニューロンネットワークの教師信号とし、学習を行っている。1つのシステムを表現しているニューロンネットワークでも、汎化能力が十分でなければ、その出力信号の特性は、既学習と未学習の各入力信号に対して異なる。このため、前章で行った学習方法では、既学習の入力信号に対しては、誤差の全くない高精度な出力信号が得られるが、未学習の入力信号に対しては、高精度な出力信号を得ることは難しいと推測される。その理由は、前段のニューロンネットワークの出力信号の特性が、既学習と未学習の各入力信号に対して異なるにもかかわらず、後段のニューロンネットワークの出力信号は、既学習の入力信号に対する前段の出力信号を補償する特性しかなく、その出力信号が、前段のニューロンネットワークの誤差の補償を反って妨げるものになるからであると考えられる。

そこで、本章の目的は、汎化利用時のECNNPの後段のニューロンネットワークの誤差の補償効果の確認、ECNNPの汎化能力の評価、予測符号化を予測器にECNNPを用いて改善を行うことにある。その結果、汎化利用時のECNNPの後段のニューロンネットワークの誤差の補償効果の有効性を確認することができた。また、予測符号化のデータ圧縮率と再構成系列信号の精度を改善できるIQPC[2]、予測符号化の予測器にECNNPを用いた場合、ECNNPの汎化能力が低くても、汎化精度を向上させることができる方法として、リアルタイムにECNNPのニューロンネ

ットワークを学習しながら予測誤差を低減するIQLPC，ニューロンネットワークの教師信号の符号にそのニューロンネットワークの出力信号の符号を合わせて誤差の低減を行う方法，IQLPCとCABを用いた2重圧縮の4つの考案に関して述べる．

## 8.2 独立量子化型学習予測符号化

### 8.2.1 独立量子化

従来の予測符号化に用いられる予測器の予測は，常時量子化誤差の影響を受ける．したがって，高精度な再構成系列信号は，量子化誤差系列信号から得ることができない．そこで，この問題を除くための高精度な予測符号化として，IQPC を提案する．図 8.1 にその原理を示す．この図で， $x_q$  は量子化原信号， $x_p$  は予測信号， $x_{pq}$  は量子化予測信号， $e_q$  は量子化誤差信号である．符号処理は，原系列信号と予測系列信号の量子化を独立に処理することを特徴とする．したがって，予測によって，量子化誤差系列信号は影響を受けない．すなわち，この予測符号化による量子化誤差は，原系列信号の量子化によるものだけとなる．

現在のほとんどの信号は，取得時に量子化されている．したがって，IQPC による原系列信号の量子化器は，しばしば不必要となる．取得時の原系列信号の量子化幅の大きさが，IQPC での原系列信号の量子化幅の大きさと等しいならば，IQPC で誤差のない原系列信号が完全に再構成される．すなわち，データの可逆圧縮が可能となる．過去の研究に，これを画像データ

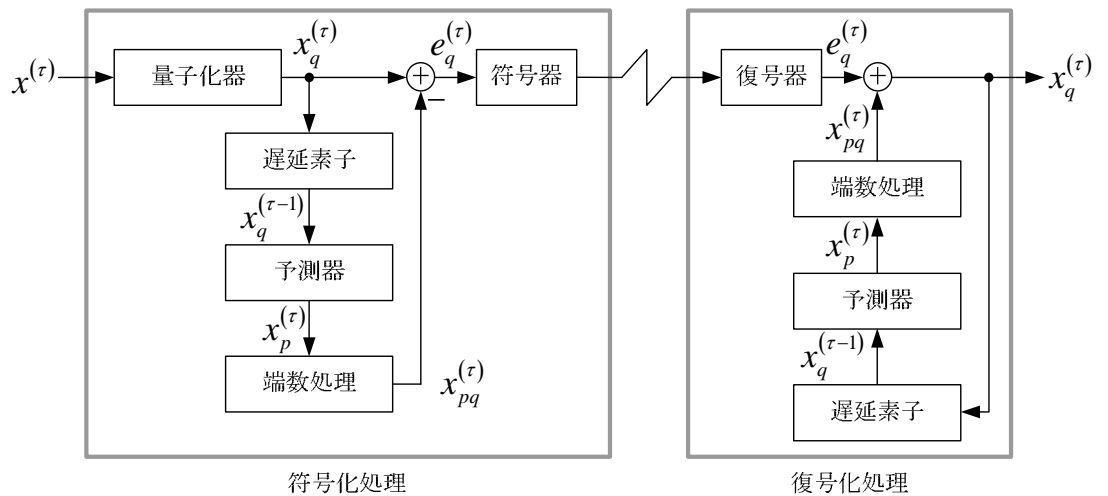


図 8.1 独立量子化型予測符号化の原理



の可逆圧縮に限定して行ったものはある[3]．また，IQPC の予測器に 1 入力の ECNNP を用いることで，データ圧縮率の改善が期待できる．

### 8.2.2 学習機能の導入

従来の予測符号化で用いられる予測器のモデルパラメータは，予測対象に従う信号特性を予め学習などによって決定し，固定されている．したがって，予測器の汎化能力が低い場合，その汎化精度は低く，予測符号化の誤差系列信号の振幅が大きくなるという問題がある．この問題を解決する方法として，予測器を学習しながら予測符号化を行う処理が考えられる．この処理を IQPC へ適用したものが IQLPC である．図 8.2 にその原理を示す．

次に IQLPC の原理を述べる．IQLPC は，符号化処理部と復号化処理部の予測器が，学習可能な同一のもので構成される．それらは，前処理として，予測対象の信号特性の学習を行っている．前処理された予測器のモデルパラメータの値は，IQLPC 開始時の初期値として設定される．符号化処理の手順は，次の通りである．

- 1) 予測器へ  $x_q^{(\tau-1)}$  を入力した後，予測器は， $x_q^{(\tau)}$  を教師信号とし，学習を行う．この学習は，終了条件を満たすまで繰り返される．この条件は，復号化処理部の予測器の学習終了条件と

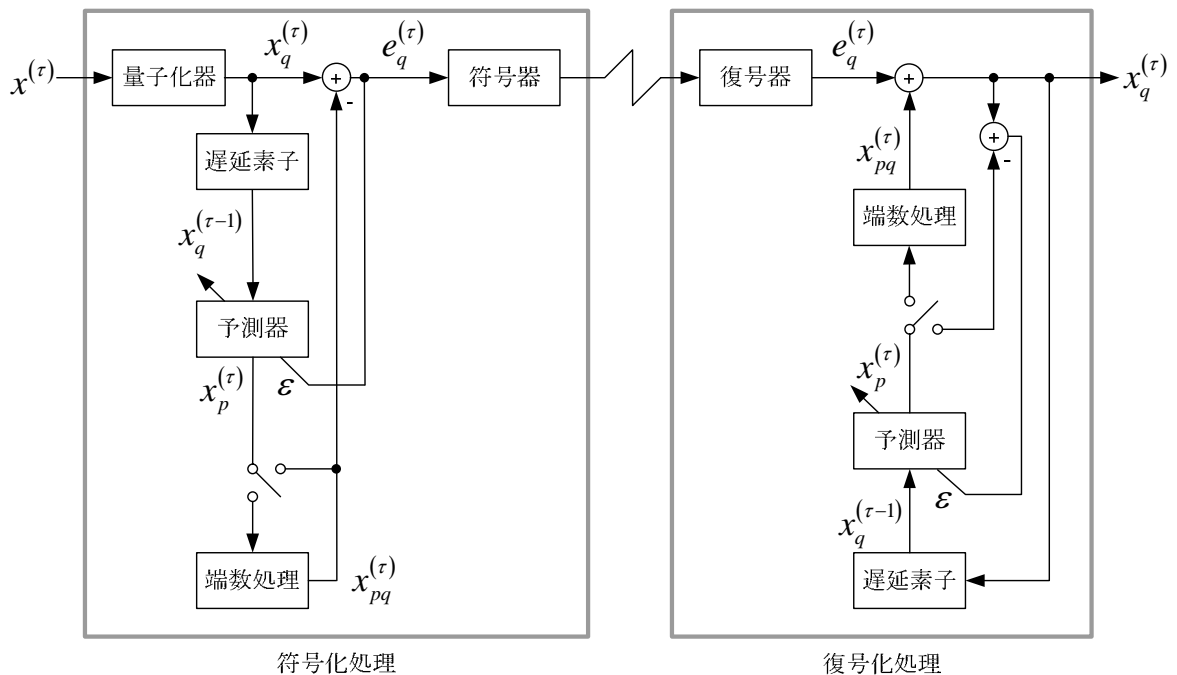


図 8.2 独立量子化型学習予測符号化の原理

しても用いる.

- 2) 学習終了後, 予測器から  $x_p^{(r)}$  が出力される.
- 3)  $x_p^{(r)}$  を端数処理し,  $x_{pq}^{(r)}$  が出力される.
- 4)  $e_q^{(r)}$  は,  $x_q^{(r)}$  から  $x_{pq}^{(r)}$  を減算して得られる.
- 5) 予測器の学習に用いられる誤差信号  $\varepsilon$  は,  $x_q^{(r)}$  から  $x_p^{(r)}$  を減算して得られる.
- 6)  $e_q^{(r)}$  は, 符号器で符号化され, 復号化処理部の復号器や記憶装置などへ伝送される.

また, 復号化処理の手順は, 次の通りである.

- 1) 符号化処理部や記憶装置などから伝送される符号化信号は, 復号器によって, 復号化される. ここで, 復号器から  $e_q^{(r)}$  が出力される.
- 2)  $e_q^{(r)}$  に  $x_{pq}^{(r)}$  を加算し,  $x_q^{(r)}$  を再構成する.
- 3) ここで,  $x_{pq}^{(r)}$  は,  $x_p^{(r)}$  を端数処理して得られている.  $x_p^{(r)}$  は, 予測器の学習終了後, 予測器から出力されている.
- 4)  $\varepsilon$  は,  $x_q^{(r)}$  から  $x_p^{(r)}$  を減算して得られる.
- 5) 予測器へ  $x_q^{(r-1)}$  を入力した後, 予測器は,  $x_q^{(r)}$  を教師信号とし, 学習が行われている. この学習は, 終了条件を満たすまで繰り返されている. この条件は, 符号化処理部の予測器の学習終了条件と同じである.

### 8.2.3 誤差の低減化

予測符号化は, その誤差系列信号の振幅を小さくさせることで, データ圧縮率を改善することができる. その方法として, ニューロンネットワークの教師信号の符号にそのニューロンネットワークの出力信号の符号を合わせることで, 誤差の低減ができることを示す. これを IQLPC に用いる ECNNP の各段のニューロンネットワークの復元後の出力に適用することで, データ圧縮率の改善が図れる. これを式(8.1), (8.2)に示す.

$$\begin{cases} \text{sgn}(y_i^{(r)} + b_i) = \text{sgn}(z_i^{(r)} + b_i) \\ y_{i+1}^{(r)} = y_i^{(r)} - z_i^{(r)} \end{cases} \quad (i=1,2,\dots,n) \quad (8.1)$$

$$\begin{cases} \text{sgn}(y_i^{(r)} + b_i) \neq \text{sgn}(z_i^{(r)} + b_i) \\ y_{i+1}^{(r)} = y_i^{(r)} + z_i^{(r)} + 2b_i \end{cases} \quad (i=1,2,\dots,n) \quad (8.2)$$

ここで、 $b_i$ は ECNNP の第  $i$  段目のニューロンネットワークの教師信号と出力信号のバイアス値である。

## 8.3 計算機シミュレーション

### 8.3.1 方法

ECNNPの汎化能力とIQLPCの評価方法を説明する．ECNNPの汎化能力を評価するために，第7章の7.3.1 a.の学習能力の改善で用いた2段のID2VNNの緒元を基に，その遅延素子数とフィルタ長を2倍にした別の数段のID2VNNで構成されるECNNPと正常洞調律心電図を用いて，第7章と同様な心電図信号予測に関する学習計算機シミュレーションを行う．学習用信号として，図8.3のMIT-BIH No.16786の正常洞調律心電図信号の開始から4,000個を用いる．

はじめに， $A_{in1}=A_1=1/3.275$ とし， $NN_1$ へ空間方向が1次元の時系列パターンを入力信号 $x(\tau)$ と教師信号 $y_1(\tau)=x(\tau+1)$ の組合せを与え探索学習を行う．ここで，最小平均RMSEを得た条件の中から最小RMSEのID2VNNを選び，それを $NN_1$ とする．さらに，この $NN_1$ の出力信号をECNNPの教師信号レベルへ復元し，その信号とECNNPの教師信号から得られた誤差信号を $NN_2$ の学習用信号の一部とする．

次に， $NN_2$ へ空間方向が2次元の時系列パターンを入力信号 $x(\tau)$ ， $x_{21}(\tau)$ と空間方向が1次元の時系列パターンの教師信号 $y_2(\tau)=x_{21}(\tau+1)$ の組合せを与え探索学習を行う．入力信号 $x(\tau)$ には，図

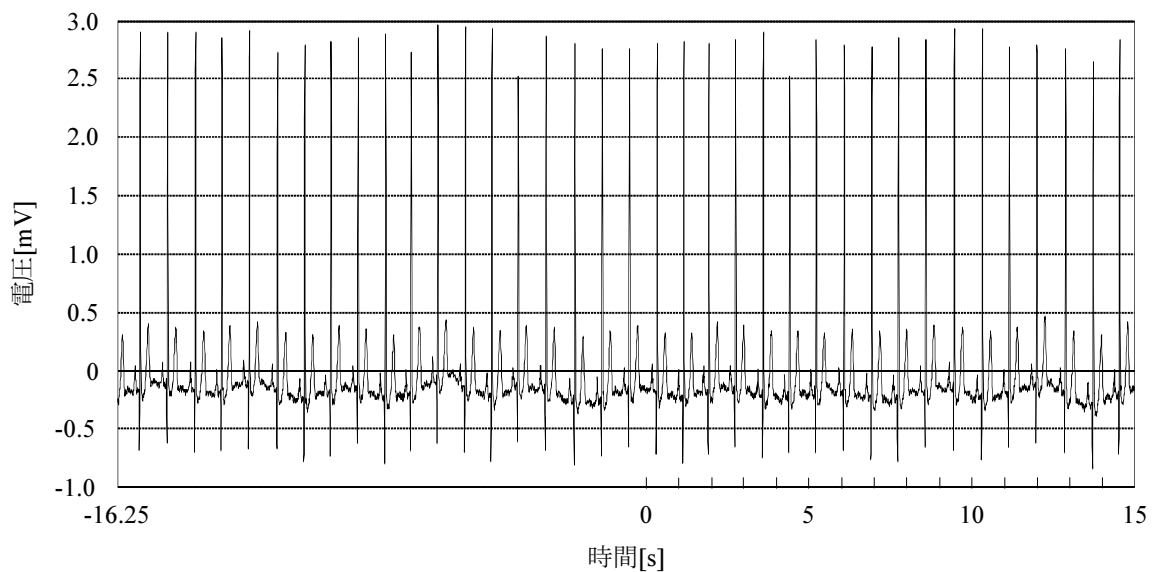


図8.3 学習用正常洞調律心電図信号

表8.1 学習計算機シミュレーション条件

段 数			第1段目	第2段目以降
学習則			2VNNの学習則	
初 期 値	結合荷重		-0.1～0.1	
	しきい値		-0.1～0.1	
	予測	$\sigma_1$	$0.7 \times 0.3^p$	
	係数	$\sigma_2$	$0.7 \times 0.3^p \times 0.7 \times 0.3^q$	
中間層素子数			4	10
フィルタ長			137	127
タップ数			283	
学習	勾配	範囲	$10^{-6} \sim 1$	
強化	法項	間隔	10倍	
係数	慣性項		0	
出力係数			1	
学習サイクル			1,000	
試行回数			3	

8.3の心電図信号，入力信号 $x_2(\tau)$ と教師信号 $y_2(\tau)$ には，第7章の式(7.4)の初期条件と $NN_2$ の学習用誤差信号から求められる信号を用いる．ここで， $Ae_{in2}=A_2$ ， $a_{21}=1/3.275$ とし， $Ae_{in2}$ ， $A_2$ には， $NN_2$ の学習用誤差信号の最大絶対値を1にする値を用いる．ここで，最小平均RMSEを得た条件の中から最小RMSEのID2VNNを選び，それを $NN_2$ とする．さらに， $NN_1$ と $NN_2$ の出力信号をECNNPの教師信号レベルへ復元し，それらの信号とECNNPの教師信号から得られた誤差信号を第3段目のニューロンネットワーク，以下 $NN_3$ と称するニューロンネットワークの学習用信号の一部とする．

$NN_2$ の決定方法と同様に，第3段目以降のニューロンネットワークを順次決定して行く．そして， $NN_1$ から最終段のニューロンネットワークまでの出力信号をECNNPの教師信号レベルへ復元し，それらの総和から得たECNNPの出力信号とECNNPの教師信号から得られる予測誤差が完全になくなった時点で，ニューロンネットワークの決定を終了する．

学習は，2,080単位時間のデータを初期入力後，入力信号と教師信号の対を1回提示して行われる．この過程を1学習サイクルとし，これを繰り返す．計算機シミュレーションは，勾配法項の学習強化係数を変化させながら，3回の探索学習を行い，その結果得られた平均RMSEを比較する．表8.1に学習計算機シミュレーション条件を示す．ここで，予測係数の初期値は，指数平滑法を用いる．その他の初期値は，学習計算機シミュレーション1回ごとに，擬似乱数で決定される．

以上の探索学習で得られたECNNPへ図8.3の心電図信号データを初期入力し，その後，その

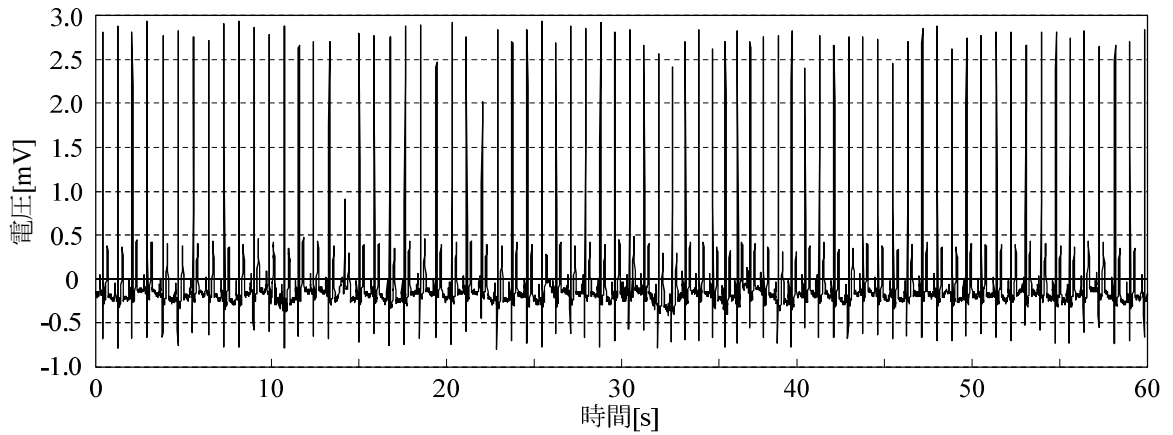


図8.4 テスト用正常洞調律心電図信号

信号と連続する図8.4の1.875秒間240個の未学習信号データを入力して、その汎化出力を評価する。

次に、IQLPCの評価方法を説明する。第7章で実現した2段のID2VNNから構成されるECNNPをIQLPCの予測器に用いて、予測符号化を評価する。IQLPCへ入力する原系列信号には、前項と同じMIT-BIH No.16786の被験者の心電図信号の学習終了直後から始まる図8.4の60秒間の未学習信号7,680個を用いる。はじめに、ECNNPの第1段目のID2VNNの学習サイクルを0回から5回まで1回ずつ増加させながらIQLPCを行う。次に、最小最大予測誤差振幅を実現した学習サイクルの第1段目のID2VNNを用いて、第2段目のID2VNNの学習サイクルを0回から5回まで1回ずつ増加させながらIQLPCを行う。ここで、通常の符号化の必要ビット数は、最大予測誤差振幅で決定されるため、最小最大予測誤差振幅を指標として用いる。以上のIQLPCより得られる最小最大予測誤差振幅を実現した各段の学習サイクルの組合せのECNNPを評価する。さらに、最良の学習サイクルの組合せのECNNPの第2段目のID2VNNへ8.2.3の誤差の低減方法をバイアス値なしで施し、最大予測誤差振幅を用いてIQLPCを評価する。最後に、その誤差を、最小値をゼロとした量子化後、符号なしでビット変換を行い、2重圧縮として、平均的に可逆圧縮性能が高いと評価されているCABを用いた圧縮を行う。それらの結果得られたデータサイズを原系列信号データや原系列信号データをCABで圧縮したデータのデータサイズと比較し、圧縮率を用いて評価する。ここで、圧縮に用いるソフトウェアのプログラムファイル容量やデータのディスク上の物理サイズに関しては、論じないこととする。

### 8.3.3 結果

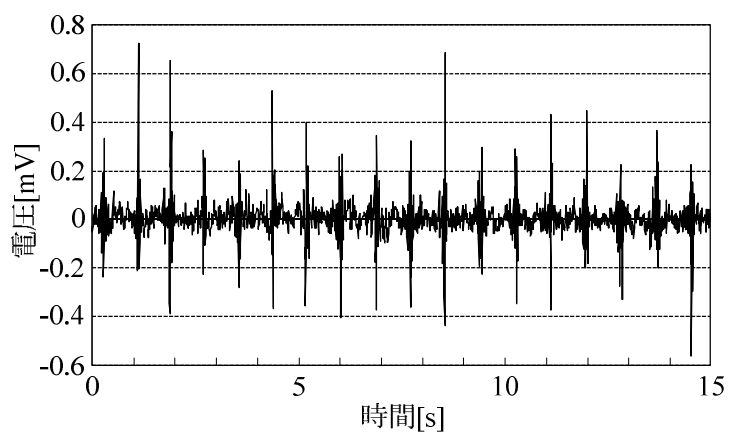
#### a. 誤差収束型予測器

ECNNP のニューロンネットワークを第 1 段目から ID2VNN の学習により順次決定して行くと、第 4 段目のニューロンネットワークを決定した時点で、ECNNP の出力誤差はなくなり、完全学習が実現した。図 8.5 に、ECNNP の第 2 段目から第 4 段目までの ID2VNN の学習用誤差信号を示す。これらの図より、誤差信号の振幅が、ニューロンネットワークの段数の増加に伴って小さくなることが確認でき、ECNNP の出力誤差が収束して行くことが示されている。図 8.6 に、第 1 段目から第 4 段目までの ID2VNN の学習強化係数に対する平均 RMSE を示す。この図からも、各段の最小平均 RMSE が、ニューロンネットワークの段数の増加に伴って、1 段当りおよそ 10 分の 1 ずつ指数関数的に減少することが確認でき、ECNNP の出力誤差が収束して行くことが示されている。図 8.7 に、第 1 段目から第 4 段目までの ID2VNN の最小平均 RMSE を実現した ID2VNN の学習回数に対する 1 単位時間データ当りの平均評価関数値の平均を示す。この図より、前段の ID2VNN の学習終了時の誤差が、後段の ID2VNN の学習開始時の誤差として引き継がれ、ニューロンネットワークの段数の増加とともに、学習が順調に進行していることが確認できる。図 8.8 に、学習後の ECNNP の出力信号を示す。これは、教師信号と完全に一致している。以上の結果より、僅か 1,000 回の ECNNP の学習でも、ニューロンネットワークの段数を増加させることで、完全学習が得られることが示された。

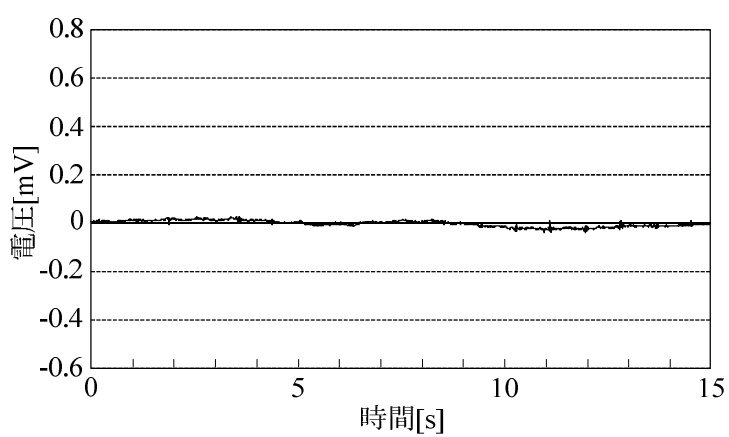
次に、学習後の ECNNP へ未学習信号データを入力して得られた ECNNP の汎化出力の結果を表 8.2 に示す。この表より、1.875 秒後の予測誤差の絶対値平均は、ニューロンネットワークの段数の増加とともに減少し、すなわち、汎化能力が向上していることが確認できる。以上より、汎化利用時の ECNNP の後段のニューロンネットワークの誤差の補償効果の有効性を確認することができた。

#### b. 独立量子化型学習予測符号化

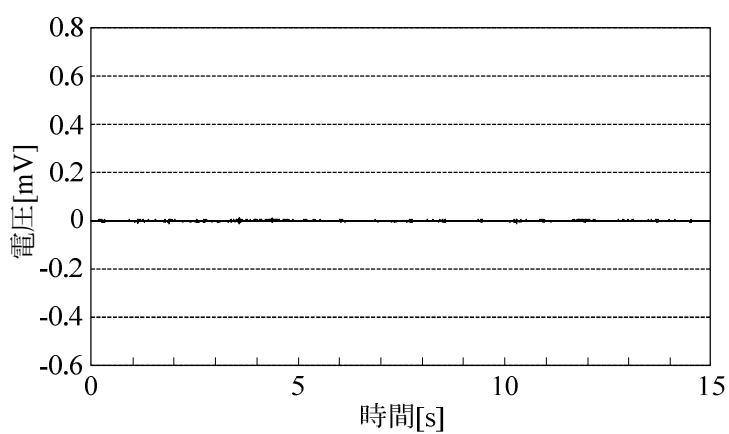
図 8.9 に、IQLPC で用いる ECNNP の第 1 段目と第 2 段目の ID2VNN の学習サイクルと最大予測誤差振幅の絶対値の関係を示す。これより、第 1 段目の ID2VNN の最大予測誤差振幅の絶対値は、学習を全く行わない場合が、最も小さいということが確認できる。また、この第 1 段目の ID2VNN を用いたとき、第 2 段目の ID2VNN の最大予測誤差振幅の絶対値は、学習サイクルが 2 回の場合が、最も小さいということが確認できる。さらに、これら最良の学習サイクルの組合せで得られた ECNNP の第 2 段目の ID2VNN へ 8.2.3 の誤差の低減方法をバイアス値なしで施した結果、第 2 段目の ID2VNN の最大予測誤差振幅の絶対値、すなわち、ECNNP の



(a) 第2段目



(b) 第3段目



(c) 第4段目

図 8.5 第2段目から第4段目までの入力遅れ2次ボルテラニューロンネットワークの学習用誤差信号

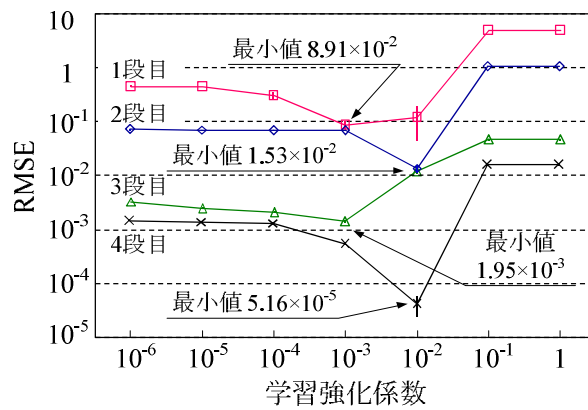


図 8.6 第 1 段階目から第 4 段階目までの入力遅れ 2 次ボルテラニューロンネットワークの学習強化係数に対する二乗平均平方根誤差の平均値

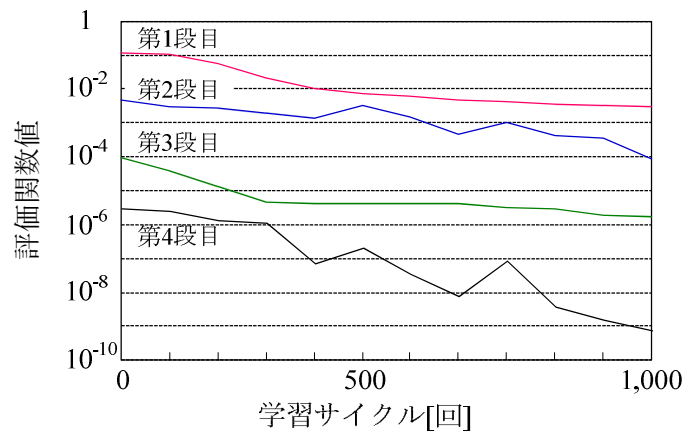


図 8.7 第 1 段階目から第 4 段階目までの二乗平均平方根誤差の最小平均値を実現した入力遅れ 2 次ボルテラニューロンネットワークの学習回数に対する 1 単位時間当りの平均評価関数値の平均値

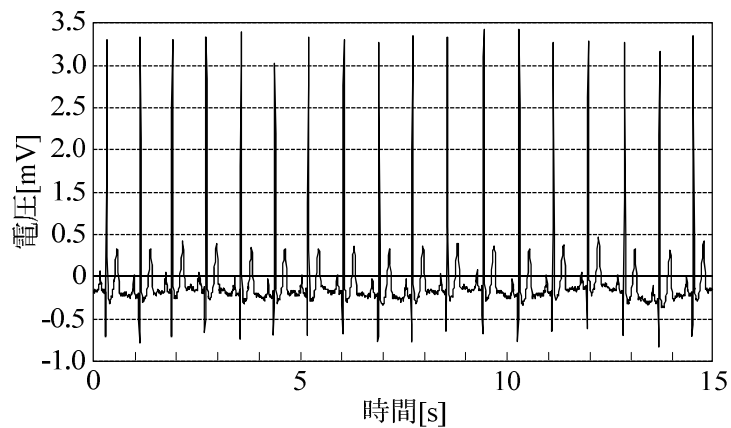


図 8.8 学習後の誤差収束型ニューロンネットワーク予測器の出力信号



表8.2 未学習信号データ入力時の誤差収束型ニューロンネットワーク予測器の各段の入力遅れ2次ボルテラニューロンネットワークまでの予測誤差の絶対値平均

段数	第1段目	第2段目	第3段目	第4段目
予測誤差の絶対値平均[mV]	$2.915 \times 10^{-1}$	$2.880 \times 10^{-1}$	$2.871 \times 10^{-1}$	$2.870 \times 10^{-1}$
百分率[%]	100.0	98.79	98.49	98.46

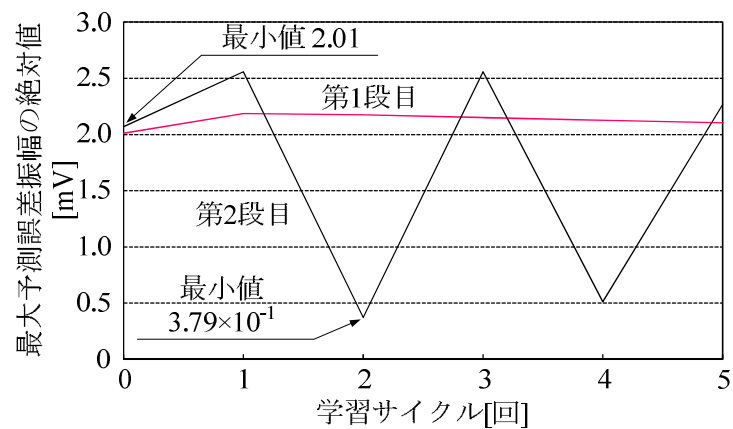


図8.9 誤差収束型ニューロンネットワーク予測器の各段の入力遅れ2次ボルテラニューロンネットワークの学習サイクルに対する最大誤差振幅の絶対値

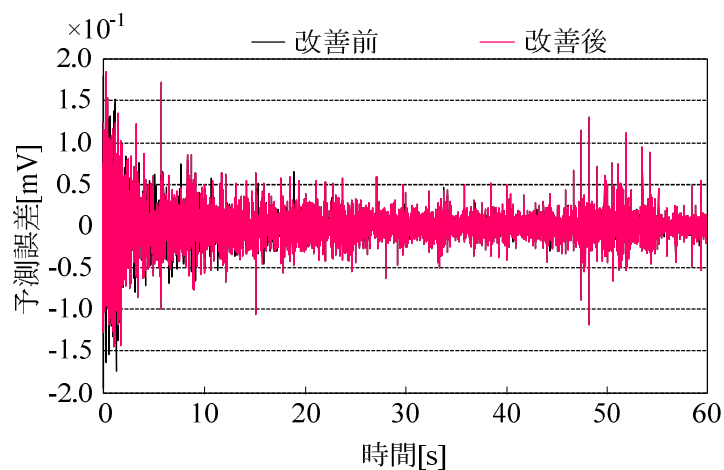


図8.10 最小最大予測誤差振幅を実現した独立量子化型学習予測符号化の未量子化の誤差系列信号

表 8.3 圧縮方法による圧縮率の比較

圧縮方法	無圧縮	IQLPC	CAB	IQLPC+CAB
データサイズ[byte]	11,520	6,720	6,676	5,712
平均ビット長[bit]	12.0	7.00	6.95	5.95
圧縮率	1	1.71	1.73	2.02

最大予測誤差振幅の絶対値は、 $3.29 \times 10^{-1} [\text{mV}]$ であり、改善前より13.2%減少した。このときの改善前後に得られたECNNPの予測誤差系列信号を図8.10に示す。これより、予測誤差は、予測開始時に大きく、時間の経過とともに徐々に小さくなるが、途中再び増大していることが確認できる。また、ECNNPの改善後は、予測開始時の予測誤差が小さくなり、最大予測誤差振幅が小さくなっていることが確認できる。

次に、各圧縮方法による圧縮率の比較結果を表8.3に示す。これより、IQLPCとCABは、ほぼ同等の圧縮能力を備えており、圧縮率は、IQLPCとCABで2重圧縮を施す方法が最も高く、無圧縮のデータサイズを50%未満にできるということが確認できる。また、この方法の圧縮率は、CABより16.8%向上している。

## 8.4 考察

はじめに、ECNNPの汎化能力について検討する。ECNNPは、学習回数が少なくても、ニューロンネットワークの段数を増加させることで、完全学習が実現できるということが実証された。これは、第6章の6.4の考察で述べたように、ECNNPに関してもニューロンネットワークの段数の増加とともに、出力の高精度化と学習の高速化の効果が十分期待できる結果であるととらえることができる。従来のニューロンネットワークの汎化能力に関する説では、教師信号に対する学習精度を高めた場合、汎化能力に支障をきたすとしていたが、今回の結果では、それが必ずしも起こるとは限らないということが示されている。この原因は、学習時と同様に、汎化時においても、後段のニューロンネットワークの誤差の補償効果の有効性が存在するということである。汎化時の誤差補償効果は、学習で得られた入力信号と教師信号の平均的な関数関係と汎化時の入力信号と出力信号の関数関係が類似している場合に得られる。すなわち、汎化能力確認のシミュレーション結果より、今回の正常洞調律心電図信号に関して、誤差補償効果が得られるための上記条件が、学習信号の直後の1.875秒間の未学習信号に含まれているため、その有効性が得られたと考えられる。さらに、ECNNP、ニューロンネットワーク、入力出

力信号などを工夫することで、汎化能力の向上が期待できる。

次に、IQLPC について検討する。予測符号化を IQLPC とし、その予測器に ECNNP を用い、ニューロンネットワークの教師信号の符号に出力信号の符号を合わせる誤差の低減方法を導入し、さらに CAB を用いて 2 重圧縮することで、CAB 単独の圧縮率より高い圧縮率の原系列信号に対する可逆圧縮が実現できた。しかし、さらに検討をしなければならないことは、ECNNP の学習終了条件である。学習終了条件としては、学習回数、予測誤差、それらの組合せ、学習用誤差信号に応じた条件などが考えられる。本章では、学習回数に関してだけの確認であるため、様々な学習終了条件の中で、データ圧縮率が最も大きくなる条件をさらに調べることで、ECNNP の汎化能力の改善ができると考えられる。また、さらにデータ圧縮率を向上させる方法として、IQLPC と他のデータ圧縮方法とを組み合わせでの使用などが考えられる。例えば、予測誤差信号系列の特徴に合わせた独自の圧縮方法や、圧縮順序を変える手法の考案などである。これらも、データ圧縮の有効性を確認しなければならない。

最後に、ニューロンネットワークの教師信号の符号に出力信号の符号を合わせる誤差の低減方法の応用について検討する。誤差の低減方法を学習中のニューロンネットワークの出力に適用することで、学習速度を改善できる可能性がある。また、符号出力用ニューロンネットワークを用意し、それに教師信号の符号を学習させ、出力用ニューロンネットワークの出力信号の符号を符号出力用ニューロンネットワークで与えることで、出力用ニューロンネットワークの汎化出力の精度を向上させることができる可能性がある。

## 8.5 結言

正常洞調律心電図予測を学習サイクル 1000 回で完全学習した後の 4 段の ID2VNN から構成される ECNNP の汎化能力に関して、ECNNP の後段のニューロンネットワークの誤差の補償効果の確認を、1.875 秒間 240 個の未学習正常洞調律心電図信号データを用いて行った結果、その予測誤差の絶対値平均は、ニューロンネットワークの段数の増加とともに減少し、汎化能力が向上して行くことが示され、学習能力に関するとときと同様に、その有効性を確認することができた。この結果は、教師信号に対する学習精度を高めた場合、汎化能力に支障をきたすとしていた従来のニューロンネットワークの汎化能力に関する説が、学習で得られた入力信号と教師信号の平均的な関数関係と汎化時の入力信号と出力信号の関数関係が類似している場合には、必ずしも成り立つとは限らないということを示していると考えられる。

次に、ECNNPの汎化能力が低くても、汎化精度を向上させることができる方法として、従来の予測符号化を検討した結果、原系列信号と予測系列信号を独立に量子化して再構成系列信号の精度を向上させるIQPC、予測器を学習しながら予測誤差を低減するIQLPC、ニューロンネットワークの教師信号の符号にそのニューロンネットワークの出力信号の符号を合わせて誤差の低減を行う方法、IQLPCとCABを用いた2重圧縮を考案した。それらの計算機シミュレーションとして、第7章で実現した2段のID2VNNから構成されるECNNPをIQLPCの予測器に用いて、60秒間の未学習正常洞調律心電図信号7,680個のデータを圧縮した結果、IQLPCは、CABと同程度の圧縮率が得られることが示された。さらに、IQLPCとCABを用いた2重圧縮では、圧縮率2.02が得られ、CAB単独の圧縮率より16.8%向上することが示された。

今後の課題は、ECNNPの汎化能力の改善、IQLPCの圧縮率を最も大きくできるECNNPの学習終了条件や圧縮率の改善方法の検討、誤差の低減方法の応用の検証などである。

## 参考文献

- [1] 小早川俊祐, 藤井崇文, 横井博一, “2 次ボルテラニューロンネットワークによる心電図の非線形予測,” バイオメディカル・ファジィ・システム学会和文誌, vol.11, no.2, pp.101-111, 2009.
- [2] S. Kobayakawa and H. Yokoi, “Proposal of Predictive Coding Using Error Convergence-type Neuron Network System,” The Processings of the ISCA 22nd International Conference on Computers and Their Applications in Industry and Engineering, San Francisco, USA, pp.169-174, 2009.
- [3] A. Fukunaga and A. Stechert, “Evolving Nonlinear Predictive Models for Lossless Image Compression with Genetic Programming,” Proceedings of the 3rd Annual Genetic Programming Conference, San Francisco, USA, pp.95-102, 1998.

## 第9章 結論

### 9.1 本論文の要約

ニューロンネットワークの基本性能を決定付ける学習精度に関する改善を行い，入力信号と教師信号との間の関数関係に対して十分に学習を行っても，出力誤差が全くない完全学習を保証できないという問題が残っていた．予測符号化では，符号化と復号化の二つの処理で，予測器が用いられている．この予測器には，非線形予測法で非線形予測を行うニューロンネットワークがあるが，前記問題点のため高精度な学習や汎化が得られないままの圧縮結果を得ていた．本研究の目的は，この学習精度に関する問題を解決し，学習精度を高めた予測器で汎化精度を向上させ，高圧縮率の量子化原系列信号に関する可逆圧縮方法の予測符号化の提案と完全学習後に得られる時系列信号の保存機能の実現であった．研究の結果，次の結論が得られた．

第2章では，2入力BPNと5入力BPNを用いた2つの非線形時系列信号予測システムの学習精度を評価し，5入力BPNを用いた非線形時系列信号予測システムの方が，2入力BPNを用いた非線形時系列信号予測システムより学習精度が向上することを示した．

第3章では，BPNとPNNを5入力4出力非線形時系列信号予測システムへ適用して学習精度を評価し，すべての出力において，PNNの学習精度の方が，BPNより高いことを示し，その原因を説明した．

第4章では，IDNNと2VNNによる正常洞調律心電図の非線形時系列信号予測学習を評価し，両ネットワークの学習精度が同等であることを示した．

第5章では，学習計算機シミュレーション時間が比較的少なく，かつ，非線形予測能力がある1VNNを用いて，TVVNNを構成し，不整脈心電図信号予測に関してのTVVNNの有効性を評価する方法を検討した．

第6章では，関数関係がある入出力を任意の近似関数と近似関数によって生じる誤差を次々と別の近似関数で表した汎関数級数の誤差を収束させる誤差収束の原理を考案した．この原理は，任意の近似関数の数を無限にすることで，数学的に汎関数級数の誤差項がゼロとなることを保証する．そして，誤差収束の原理を実際の工学処理に発展させた量子化信号の誤差収束の原理とそれを適用したECFSも考案した．特に，任意の近似関数にニューロンネットワークを用いたものをECNNと定義した．また，ECFSを並列ユニットとしたECPFSも考案した．さらに，ECFSとECPFSを用いた応用として，非線形プラントのシミュレータや系列信号逐次出力器として使えることが期待できるECRSとECPRSも考案した．

第7章では、ECNNの各段のニューロンネットワークの入力信号と教師信号の間の関数関係がほとんどない場合の学習困難性を、ECNNの各段のニューロンネットワークを予測器として構成し、その入力信号と教師信号の間の関数関係を強くすることで改善したECNNPを考案した。次に、正常洞調律心電図信号予測学習に対して、2段の2VNNから構成されるECNNPの学習計算機シミュレーションを行い、完全学習能力がECNNPにあることを実証した。さらに、2VNNで構成されるECNNPの学習能力をID2VNNで構成されるECNNPで改善できることを示した。

第8章では、ECNNPのニューロンネットワークの段数の増加とともに汎化能力が向上して行くことが示された。また、予測符号化のデータ圧縮率と再構成系列信号の精度を改善できるIQLPC、リアルタイムにECNNPのニューロンネットワークを学習しながら予測誤差を低減させるIQLPC、ニューロンネットワークの教師信号の符号にそのニューロンネットワークの出力信号の符号を合わせて誤差の低減を行う方法、IQLPCとCABを用いた2重圧縮の4つの考案を行った。その結果、CABを超える圧縮率を示した。

## 9.2 今後の課題

ECNNPは、各段のニューロンネットワークの逐次学習による正常洞調律心電図の非線形性の強い時系列信号の予測学習に対しては、完全学習能力が示されたものの、同時学習による学習能力の評価はされていない。汎化能力に関する評価も行われたが、入力信号と出力信号の変化する関数関係に対応するための改善が行われていない。ECNNPを用いたIQLPCの有効性の確認は行われたが、IQLPCの最適な圧縮率を得る条件が十分調査されていない。

今後の課題としては、前記3項目の研究を行うことである。具体的には次の通りである。

1. ECNNPの各段のニューロンネットワークの同時学習による学習能力の評価。
2. 学習サイクルを少なくできるECNNPのニューロンネットワークの段数の調査。
3. 非線形プラントの複数の時系列信号に対してのECNNPの汎化能力の評価。
4. IQLPCのデータ圧縮率を最も大きくできるECNNPの学習終了条件の調査。
5. IQLPCと他のデータ圧縮方法とを組み合わせたデータ圧縮方法の有効性の確認。
6. 学習則や汎化出力決定に対する誤差の低減方法の応用の評価。

## 謝辞

本研究を進めるにあたり，博士課程前期・後期と9年間以上もの長い間，根気よく，厳しく専門の奥義や研究の精髓をご指導賜りました九州工業大学大学院生命体工学研究科の横井博一教授に深く感謝を申し上げます。また，本論文の作成や審査に際し，貴重なお時間を賜り，ゼミなどでプレゼンテーションや論文作法などのご指導をいただきました猪平栄一講師，表現に対する細心の注意を促していただきました花本剛士教授，数学的な議論を深めていただきました永松正博教授，研究の締めを励ましていただきました芹川聖一教授に深く感謝を申し上げます。また，同じ研究室でお会いし，自由かつたつな学究生活をともに過ごしていただきましたすべての学生諸氏，研究活動を支えていただきました九州工業大学大学院生命体工学研究科の職員の皆様をはじめとする大学関係諸氏に深く感謝を申し上げます。



# 本論文に関連する研究業績リスト

## 査読付学術論文

1. S. Kobayakawa and H. Yokoi, "Predictor Using an Error-Convergence Neuron Network and its Application to Electrocardiograms," Journal of Advanced Computational Intelligence and Intelligent Informatics, in contribution, pp. 1-14, 2010.
2. S. Kobayakawa and H. Yokoi, "Experimental Study of Learning Capability for Parallel-type Neuron Network," International Journal of Biomedical Soft Computing and Human Sciences, vol.16, in print, no.2, pp.1-9, 2010.
3. S. Kobayakawa and H. Yokoi, "Electrocardiogram Prediction Using Error Convergence-type Neuron Network System," International Journal of Computer and Network Security, vol.2, no.2, pp.14-23, 2010.
4. 小早川俊祐, 藤井崇文, 横井博一, "2次ボルテラニューロンネットワークによる心電図の非線形予測," バイオメディカル・ファジィ・システム学会和文誌, vol.11, no.2, pp.101-111, 2009.
5. S. Kobayakawa and H. Yokoi, "Evaluation of Prediction Capability of Non-recursion Type 2<sup>nd</sup>-order Volterra Neuron Network for Electrocardiogram," Advances in Neuro-Information Processing, 15th International Conference, ICONIP 2008, Auckland, New Zealand, November 25-28, 2008, Revised Selected Papers, Part II, Lecture Notes in Computer Science, vol.5507, pp.679-686, 2009.

## 国際会議における発表

1. S. Kobayakawa and H. Yokoi, "Proposal of Independent Quantization Predictive Coding as Learning," International Conference of the Institute for Environment, Engineering, Economics and Applied Mathematics, Proceedings, in review, Sliema, Malta, pp.1-8, 2010.
2. S. Kobayakawa and H. Yokoi, "Proposal of Error Convergence-type Predictor Using 2<sup>nd</sup>-order Volterra Neuron Networks with Input-delay," International Symposium on Application of Biomechanical Control Systems to Precision Engineering, Abstracts, Fukushima, Japan, pp.114-121,

- 2010.
3. S. Kobayakawa and H. Yokoi, "Proposal of Predictive Coding Using Error Convergence-type Neuron Network System," The Proceedings of the ISCA 22<sup>nd</sup> International Conference on Computers and Their Applications in Industry and Engineering, San Francisco, USA, pp.169-174, 2009.
  4. S. Kobayakawa and H. Yokoi, "Evaluation for Prediction Accuracies of Parallel-type Neuron Network," International MultiConference of Engineers and Computer Scientists 2009, IMECS 2009 Proceedings, vol.I, Hong Kong, China, pp.156-161, 2009.
  5. S. Kobayakawa and H. Yokoi, "Proposal of Error Convergence-type Neuron Network System," Presented Proceeding to 2008 International Symposium on Intelligent Informatics International, Kumamoto, Japan, pp.1-10, 2008.
  6. S. Kobayakawa and H. Yokoi, "Proposal of Error Convergence-type Neuron Network System," 2008 International Symposium on Intelligent Informatics, Extended Abstracts, Kumamoto, Japan, p.19, 2008.
  7. S. Kobayakawa and H. Yokoi, "Evaluation of Prediction Capability of Non-recursion Type 2<sup>nd</sup>-order Volterra Neuron Network for Electrocardiogram," ICONIP 2008 - 15th International Conference on Neural Information Processing of the Asia-Pacific Neural Network Assembly, Neural Information Processing, Auckland, New Zealand, pp.499-500, 2008.
  8. S. Kobayakawa, T. Fujii and H. Yokoi, "Evaluation of Prediction Capabilities of Neuron Networks Used for Electrocardiogram," The 5th International Symposium on Management Engineering, Proceedings, Kitakyushu, Japan, pp.156-161, 2008.
  9. S. Kobayakawa and H. Yokoi, "Evaluation of the Learning Capability of a Parallel-type Neuron Network," The First International Symposium on Information and Computer Elements 2007, Proceedings of ISICE 2007, Kitakyushu, Japan, pp.43-47, 2007.

## 国内学会等における発表

1. 小早川俊祐, 横井博一, "誤差収束型ニューロンネットワークシステムを用いた心電図予測," 情報処理学会創立 50 周年記念(第 72 回)全国大会講演論文集, 第 2 分冊, pp.99-100, 2010.
2. 小早川俊祐, 横井博一, "誤差収束型ニューロンネットワークシステムの収束能力に関する

- 基礎的実験,” 第 8 回情報科学技術フォーラム講演論文集, 2 分冊, pp.531-537, 2009.
3. 小早川俊祐, 横井博一, “並列型ニューロンネットワークの予測出力精度優位性の実験的検討,” 電子情報通信学会技術研究報告, NC2008-1~10, vol.108, no.54, pp.29-34, 2008.
  4. 小早川俊祐, 横井博一, “入力信号数に対する BP ネットワークの学習能力の評価,” 電子情報通信学会技術研究報告, SANE2007-102~124, vol.107, no.442, pp.83-86, 2008.
  5. 小早川俊祐, 藤井崇文, 横井博一, “心電図に対するボルテラニューロンネットワークの予測能力の評価,” 第 9 回日本知能情報ファジィ学会九州支部学術講演会予稿集, pp.19-22, 2007.
  6. 小早川俊祐, 藤井崇文, 横井博一, “心電図に対するニューロンネットワークの非線形予測能力の評価,” バイオメディカル・ファジィ・システム学会第 20 回年次大会講演論文集, pp.9-12, 2007.
  7. 小早川俊祐, 横井博一, “並列化ニューロンネットワークの航空機における予測問題への適用,” 電子情報通信学会技術研究報告, SANE2006-119~133, vol.106, no.471, pp.43-45, 2007.
  8. 小早川俊祐, 横井博一, “並列化ニューロンネットワークの予測能力評価,” 第 8 回日本知能情報ファジィ学会九州支部学術講演会予稿集, pp.3-6, 2006.
  9. 小早川俊祐, 横井博一, “航空機ピッチ姿勢制御のためのボルテラフィルタ内蔵型ニューラルネットワーク,” 平成 17 年度電気関係学会九州支部連合大会講演論文集,” p.429, 2005.